



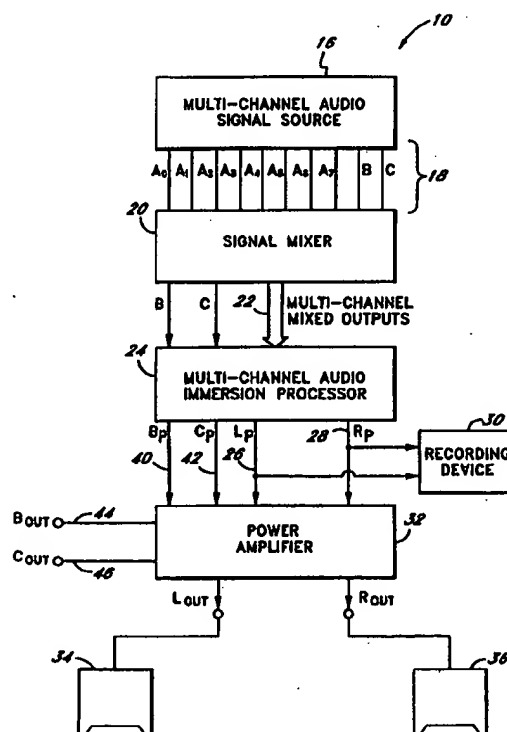
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : H04S 3/00, 1/00		A1	(11) International Publication Number: <b>WO 98/20709</b>
			(43) International Publication Date: 14 May 1998 (14.05.98)
(21) International Application Number: PCT/US97/19825 (22) International Filing Date: 31 October 1997 (31.10.97) (30) Priority Data: 08/743,776                      7 November 1996 (07.11.96)      US (71) Applicant: SRS LABS, INC. [US/US]; 2909 Daimler Street, Santa Ana, CA 92705 (US). (72) Inventors: KLAYMAN, Arnold, I.; 16821 Phelps Lane, Huntington Beach, CA 92649 (US). KRAEMER, Alan, D.; 17661 Shadel Drive, Tustin, CA 92680 (US). (74) Agent: SIMPSON, Andrew, H.; Knobbe, Martens, Olson and Bear, 16th floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>

(54) Title: MULTI-CHANNEL AUDIO ENHANCEMENT SYSTEM FOR USE IN RECORDING AND PLAYBACK AND METHODS FOR PROVIDING SAME

## (57) Abstract

An audio enhancement system and method (10) for use receives a group of multi-channel audio signals (18) and provides a simulated surround sound environment through playback of only two output signals (26 and 28). The multi-channel audio signals (18) comprise a pair of front signals intended for playback from a forward sound stage and a pair of rear signals intended for playback from a rear sound stage. The front and rear signals are modified in pairs by a multi-channel audio immersion processor (24). The multi-channel audio immersion processor (24) separates an ambient component of each pair of signals from a direct component and processing at least some of the components with a head-related transfer function. Processing of the individual audio signal components is determined by an intended playback position of the corresponding original audio signals. The individual audio signal components are then selectively combined with the original audio signals to form two enhanced output signals L<sub>OUT</sub> and R<sub>OUT</sub> for generating a surround sound experience upon playback.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**MULTI-CHANNEL AUDIO ENHANCEMENT SYSTEM  
FOR USE IN RECORDING AND PLAYBACK  
AND METHODS FOR PROVIDING SAME**

**Field of the Invention**

5           This invention relates generally to audio enhancement systems and methods for improving the realism and dramatic effects obtainable from two channel sound reproduction. More particularly, this invention relates to apparatus and methods for enhancing multiple audio signals and mixing these audio signals into a two channel format for reproduction in a conventional playback system.

**Background of the Invention**

10           Audio recording and playback systems can be characterized by the number of individual channel or tracks used to input and/or play back a group of sounds. In a basic stereo recording system, two channels each connected to a microphone may be used to record sounds detected from the distinct microphone locations. Upon playback, the sounds recording by the two channels are typically reproduced through a pair of loudspeakers, with one loudspeaker reproducing an individual channel. Providing two separate audio channels for recording permits individual processing  
15 of these channels to achieve an intended effect upon playback. Similarly, providing more discrete audio channels allows more freedom in isolating certain sounds to enable the separate processing of these sounds.

          Professional audio studios use multiple channel recordings systems which can isolate and process numerous individual sounds. However, since many conventional audio reproduction devices are delivered in traditional stereo, use of a multi-channel system to record sounds requires that the sounds be "mixed" down to only two individual  
20 signals. In the professional audio recording world, studios employ such mixing methods since individual instruments and vocals of a given audio work may be initially recorded on separate tracks, but must be replayed in a stereo format found in conventional stereo systems. Professional systems may use 48 or more separate audio channels which are processed individually before recorded onto two stereo tracks.

          In multi-channel playback systems, i.e., defined herein as systems having more than two individual audio  
25 channels, each sound recorded from an individual channel may be separately processed and played through a corresponding speaker or speakers. Thus, sounds which are recorded from, or intended to be placed at, multiple locations about a listener, can be realistically reproduced through a dedicated speaker placed at the appropriate location. Such systems have found particular use in theaters and other audio-visual environments where a captive and fixed audience experiences both an audio and visual presentation. These systems, which include Dolby  
30 Laboratories' "Dolby Digital" system; the Digital Theater System (DTS); and Sony's Dynamic Digital Sound (SDDS), are all designed to initially record and then reproduce multi-channel sounds to provide a surround listening experience.

          In the personal computer and home theater arena, recorded media is being standardized so that multiple channels, in addition to the two conventional stereo channels, are stored on such recorded media. One such standard is Dolby's AC-3 multi-channel encoding standard which provides six separate audio signals. In the Dolby AC-3  
35 system, two audio channels are intended for playback on forward left and right speakers, two channels are reproduced on rear left and right speakers, one channel is used for a forward center dialogue speaker, and one

channel is used for low-frequency and effects signals. Audio playback systems which can accommodate the reproduction of all these six channels do not require that the signals be mixed into a two channel format. However, many playback systems, including today's typical personal computer and tomorrow's personal computer/television, may have only two channel playback capability (excluding center and subwoofer channels). Accordingly, the information present in additional audio signals, apart from that of the conventional stereo signals, like those found in an AC-3 recording, must either be electronically discarded or mixed into a two channel format.

There are various techniques and methods for mixing multi-channel signals into a two channel format. A simple mixing method may be to simply combine all of the signals into a two-channel format while adjusting only the relative gains of the mixed signals. Other techniques may apply frequency shaping, amplitude adjustments, time delays or phase shifts, or some combination of all of these, to an individual audio signal during the final mixing process. The particular technique or techniques used may depend on the format and content of the individual audio signals as well as the intended use of the final two channel mix.

For example, U.S. Patent No. 4,393,270 issued to van den Berg discloses a method of processing electrical signals by modulating each individual signal corresponding to a preselected direction of perception which may compensate for placement of a loudspeaker. A separate multi-channel processing system is disclosed in U.S. Patent No. 5,438,623 issued to Begault. In Begault, individual audio signals are divided into two signals which are each delayed and filtered according to a head related transfer function (HRTF) for the left and right ears. The resultant signals are then combined to generate left and right output signals intended for playback through a set of headphones.

The techniques found in the prior art, including those found in the professional recording arena, do not provide an effective method for mixing multi-channel signals into a two channel format to achieve a realistic audio reproduction through a limited number of discrete channels. As a result, much of the ambiance information which provides an immersive sense of sound perception may be lost or masked in the final mixed recording. Despite numerous previous methods of processing multi-channel audio signals to achieve a realistic experience through conventional two channel playback, there is much room for improvement to achieve the goal of a realistic listening experience.

Accordingly, it is an object of the present invention to provide an improved method of mixing multi-channel audio signals which can be used in all aspects of recording and playback to provide an improved and realistic listening experience. It is an object of the present invention to provide an improved system and method for mastering professional audio recordings intended for playback on a conventional stereo system. It is also an object of the present invention to provide a system and method to process multi-channel audio signals extracted from an audio-visual recording to provide an immersive listening experience when reproduced through a limited number of audio channels.

For example, personal computers and video players are emerging with the capability to record and reproduce digital video disks (DVD) having six or more discrete audio channels. However, since many such computers and video players do not have more than two audio playback channels (and possibly one sub-woofer channel), they cannot use

the full amount of discrete audio channels as intended in a surround environment. Thus, there is a need in the art for a computer and other video delivery system which can effectively use all of the audio information available in such systems and provide a two channel listening experience which rivals multi-channel playback systems. The present invention fulfills this need.

5

#### Summary of the Invention

An audio enhancement system and method is disclosed for processing a group of audio signals, representing sounds existing in a 360 degree sound field, and combining the group of audio signals to create a pair of signals which can accurately represent the 360 degree sound field when played through a pair of speakers. The audio enhancement system can be used as a professional recording system or in personal computers and other home audio systems which include a limited amount of audio reproduction channels.

10

In a preferred embodiment for use in a home audio reproduction system having stereo playback capability, a multi-channel recording provides multiple discrete audio signals consisting of at least a pair of left and right signals, a pair of surround signals, and a center channel signal. The home audio system is configured with speakers for reproducing two channels from a forward sound stage. The left and right signals and the surround signals are first processed and then mixed together to provide a pair of output signals for playback through the speakers. In particular, the left and right signals from the recording are processed collectively to provide a pair of spatially-corrected left and right signals to enhance sounds perceived by a listener as emanating from a forward sound stage.

15

The surround signals are collectively processed by first isolating the ambient and monophonic components of the surround signals. The ambient and monophonic components of the surround signals are modified to achieve a desired spatial effect and to separately correct for positioning of the playback speakers. When the surround signals are played through forward speakers as part of the composite output signals, the listener perceives the surround sounds as emanating from across the entire rear sound stage. Finally, the center signal may also be processed and mixed with the left, right and surround signals, or may be directed to a center channel speaker of the home reproduction system if one is present.

20

According to one aspect of the invention, a system processes at least four discrete audio signals including main left and right signals containing audio information intended for playback from a front sound stage, and surround left and right signals containing audio information intended for playback from a rear sound stage. The system generates a pair of left and right output signals for reproduction from the front sound stage to create the perception of a three dimensional sound image without the need for actual speakers placed in the rear sound stage.

25

The system comprises a first electronic audio enhancer which receives the main left and right signals. The first audio enhancer processes an ambient component of the main left and right signals to create the perception of a broadened sound image across the front sound stage when the left and right output signals are reproduced by a pair of speakers positioned within the front sound stage.

30

A second electronic audio enhancer receives the surround left and right signals. The second audio enhancer processes an ambient component of the surround left and right signals to create the perception of an acoustic sound

35

image across the rear sound stage when the left and right output signals are reproduced by the pair of speakers positioned within the front sound stage.

5 A third electronic audio enhancer which receives the surround left and right signals. The third audio enhancer processes a monophonic component of the surround left and right signals to create the perception of an acoustic sound image at a center location of the rear sound stage when the left and right output signals are reproduced by the pair of speakers positioned within the front sound stage.

10 A signal mixer which generates the left and right output signals from the at least four discrete audio signals by combining the processed ambient component from the main left and right signals, the processed ambient component for the surround left and right signals, and the processed monophonic component from the surround left and right signals, wherein the ambient components of the main and surround signals are included in the left and right output signals in an out-of-phase relationship with respect to each other.

15 In another embodiment, the at least four discrete audio signals comprise a center channel signal containing audio information intended for playback by a front sound stage center speaker, and the center channel signal is combined by the signal mixer as part of the left and right output signals. In yet another embodiment, the at least four discrete audio signals comprise a center channel signal containing audio information intended for playback by a center speaker located within the front sound stage, and the center channel signal is combined with a monophonic component of the main left and right signals by the signal mixer to generate the left and right output signals.

20 In another embodiment, the at least four discrete audio signals comprise a center channel signal having center stage audio information which is acoustically reproduced by a dedicated center channel speaker. In yet another embodiment, the first, second, and third electronic audio enhancers apply an HRTF-based transfer function to a respective one of the discrete audio signals for creating an apparent sound image corresponding to the discrete audio signals when the left and right output signals are acoustically reproduced.

25 In another embodiment, the first audio enhancer equalizes the ambient component of the main left and right signals by boosting the ambient component below approximately 1 kHz and above approximately 2 kHz relative to frequencies between approximately 1 and 2 kHz. In yet another embodiment, the peak gain applied to boost the ambient component, relative to the gain applied to the ambient component between approximately 1 and 2 kHz, is approximately 8 dB.

30 In another embodiment, the second and third audio enhancers equalize the ambient and monophonic components of the surround left and right signals by boosting the ambient and monophonic components below approximately 1 kHz and above approximately 2 kHz, relative to frequencies between approximately 1 and 2 kHz. In yet another embodiment, the peak gain applied to boost the ambient and monophonic components of the surround left and right signals, relative to the gain applied to the ambient and monophonic components between approximately 1 and 2 kHz, is approximately 18 dB.

35 In another embodiment, the first, second, and third electronic audio enhancers are formed upon a semiconductor substrate. In yet another embodiment, the first, second, and third electronic audio enhancers are implemented in software.

5

According to another aspect of the invention, a multi-channel recording and playback apparatus receives a plurality of individual audio signals and processes the plurality of audio signals to provide first and second enhanced audio output signals for achieving an immersive sound experience upon playback of the output signals. The multi-channel recording apparatus comprises a plurality of parallel audio signal processing devices for modifying the signal content of the individual audio signals wherein each parallel audio signal processing device comprises.

A circuit receives two of the individual audio signals and isolates an ambient component of the two audio signals from a monophonic component of the two audio signals. A positional processing means which is capable of electronically applying a head related transfer function to each of the ambient and monophonic components of the two audio signals to generate processed ambient and monophonic components. The head related transfer functions corresponding to a desired spatial location with respect to a listener.

A multi-channel circuit mixer combines the processed monophonic components and ambient components generated by the plurality of positional processing means to generate the enhanced audio output signals. The processed ambient components are then combined in an out-of-phase relationship with respect to the first and second output signals.

In another embodiment, each of the plurality of positional processing means further includes a circuit capable of individually modifying the two audio signals and wherein the multi-channel mixer further combines the two modified signals from the plurality of positional processing means with the respective ambient and monophonic components to generate the audio output signals. In another embodiment, the circuit capable of individually modifying the two audio signals electronically applies, a head related transfer function to the two audio signals.

In another embodiment, the circuit capable of individually modifying the two audio signals electronically, applies a time delay to one of the two audio signals. In yet another embodiment, the two audio signals comprise audio information corresponding to a left front location and a right front location with respect to a listener. In still another embodiment, the two audio signals comprise audio information corresponding to a left rear location and a right rear location with respect to a listener.

In another embodiment, the plurality of parallel processing devices comprise first and second processing devices. The first processing device applies a head related transfer function to a first pair of the audio signals for achieving a first perceived direction for the first pair of audio signals when the output signals are reproduced. The second processing device applies a head related transfer function to a second pair of the audio signals for achieving a second perceived direction for the second pair of audio signals when the output signals are reproduced.

In another embodiment, the plurality of parallel audio processing devices and the multi-channel circuit mixer are implemented in a digital signal processing device of the multi-channel recording and playback apparatus.

According to another aspect of the invention, an audio enhancement system processes a plurality of audio source signals to create a pair of stereo output signals for generating a three dimensional sound field when the pair of stereo output signals are reproduced by a pair of loudspeakers. The audio enhancement system comprises a first processing circuit in communication with a first pair of the audio source signals. The first processing circuit is configured to isolate a first ambient component and a first monophonic component from the first pair of audio

signals. The first processing circuit is further configured to modify the first ambient component and the first monophonic component to create a first acoustic image such that the first acoustic image is perceived by a listener as emanating from a first location.

5 A second processing circuit which is in communication with a second pair of audio source signals. The second processing circuit is configured to isolate a second ambient component and a second monophonic component from the second pair of audio signals. The second processing circuit is further configured to modify the second ambient component and the second monophonic component to create a second acoustic image, such that the second acoustic image is perceived by the listener as emanating from a second location.

10 A mixing circuit which is in communication with the first processing circuit and the second processing circuit. The mixing circuit is configured to combine the first and second modified monophonic components in phase and combine the first and second modified ambient components out of phase to generate a pair of stereo output signals.

15 In another embodiment, the first processing circuit is further configured to modify a plurality of frequency components in the first ambient component with a first transfer function. In another embodiment, the first transfer function is further configured to emphasize a portion of the low frequency components in the first ambient component relative to other frequency components in the first ambient component. In yet another embodiment, the first transfer function is configured to emphasize a portion of the high frequency components of the first ambient component relative to other frequency components in the first ambient component.

20 In another embodiment, the second processing circuit is configured to modify a plurality of frequency components in the second ambient component with a second transfer function. In yet another embodiment, the second transfer function is configured to modify the frequency components in the second ambient component in a different manner than the first transfer function modifies the frequency components in the first ambient component.

25 In another embodiment, the second transfer function is configured to deemphasize a portion of the frequency components above approximately 11.5 kHz relative to other frequency components in the second ambient component. In yet another embodiment, the second transfer function is configured to deemphasize a portion of the frequency components between approximately 125 Hz and approximately 2.5 kHz relative to other frequency components in the second ambient component. In yet another embodiment, the second transfer function is configured to increase a portion of the frequency components between approximately 2.5 kHz and approximately 11.5 kHz relative to other frequency components in the second ambient component.

30 According to another aspect of the invention, a multi-track audio processor receives a plurality of separate audio signals as part of a composite audio source. The plurality of audio signals comprise at least two distinct audio signal pairs which contain audio information which is desirably interpreted by a listener as emanating from distinct locations within a sound listening environment.

35 The multi-track audio processor comprises a first electronic means which receives a first pair of the audio signals. The first electronic means separately applies a head related transfer function to an ambient component of



the first pair of audio signals to create a first acoustic image wherein the first acoustic image is perceived by a listener as emanating from a first location.

A second electronic means which receives a second pair of the audio signals. The second electronic means separately applies a head related transfer function to an ambient component and a monophonic component of the second pair of audio signals to create a second acoustic image wherein the second acoustic image is perceived by the listener as emanating from a second location.

A means which mixes the components of the first and second pair of audio signals received from the first and second electronic means. The means for mixing combines the ambient components out of phase to generate the pair of stereo output signals.

According to another aspect of the invention, an entertainment system has two main audio reproduction channels for reproducing an audio-visual recording to a user. The audio-visual recording comprises five discrete audio signals including a front-left signal,  $F_L$ , a front-right signal,  $F_R$ , a rear-left signal,  $R_L$ , a rear-right signal,  $R_R$ , and a center signal,  $C$ , and wherein the entertainment system achieves a surround sound experience for the user from the two main audio channels. The entertainment system comprising an audio-visual playback device for extracting the five discrete audio signals from the audio-visual recording.

An audio processing device receives the five discrete audio signals and generates the two main audio reproduction channels. The audio processing device comprises a first processor for equalizing an ambient component of the front signals,  $F_L$  and  $F_R$ , to obtain a spatially-corrected ambient component  $(F_L - F_R)_p$ . A second processor equalizes an ambient component of the rear signals,  $R_L$  and  $R_R$ , to obtain a spatially-corrected ambient component  $(R_L - R_R)_p$ . A third processor equalizes a direct-field component of the rear signals,  $R_L$  and  $R_R$ , to obtain a spatially-corrected direct-field component  $(R_L + R_R)_p$ .

A left mixer generates a left output signal. The left mixer combines the spatially-corrected ambient component,  $(F_L - F_R)_p$ , with the spatially-corrected ambient component,  $(R_L - R_R)_p$ , and the spatially-corrected direct-field component,  $(R_L + R_R)_p$ , to create the left output signal.

A right mixer generates a right output signal. The right mixer combines an inverted spatially-corrected ambient component,  $(F_R - F_L)_p$ , with an inverted spatially-corrected ambient component,  $(R_R - R_L)_p$ , and the spatially-corrected direct-field component,  $(R_L + R_R)_p$ , to create the right output signal.

A means reproduces the left and right output signals through the two main channels in connection with playback of the audio-visual recording to create a surround sound experience for the user.

In another embodiment, the center signal is input by the left mixer and combined as part of the left output signal and the center signal is combined by the right mixer and combined as part of the right output signal. In yet another embodiment, the center signal and a direct field component of the front signals,  $F_L + F_R$ , are combined by the left and right mixers as part of the left and right output signals, respectively. In still another embodiment, the center signal is provided as a third output signal for reproduction by a center channel speaker of the entertainment system.

In another embodiment, the entertainment system is a personal computer and the audio-visual playback device is a digital versatile disk (DVD) player. In yet another embodiment, the entertainment system is a television

and the audio-visual playback device is an associated digital versatile disk (DVD) player connected to the television system.

In another embodiment, the first, second, and third processors emphasize a low and high range of frequencies relative to a mid-range of frequencies. In yet another embodiment, the audio processing device is implemented as an analog circuit formed upon a semiconductor substrate. In still another embodiment, the audio processing device is implemented in a software format, the software format executed by a microprocessor of the entertainment system.

According to another aspect of the invention, a method enhances a group of audio source signals wherein the audio source signals are designated for speakers placed around a listener to create left and right output signals for acoustic reproduction by a pair of speakers in order to simulate a surround sound environment. The audio source signals comprise a left-front signal ( $L_f$ ), a right-front signal ( $R_f$ ), a left-rear signal ( $L_R$ ), and a right-rear signal ( $R_R$ ).

The method comprises an act of modifying the audio source signals to create processed audio signals based on the audio content of selected pairs of the source signals. The processed audio signals are defined in accordance with the following equations:

$$\begin{aligned} P_1 &= F_1(L_f - R_f), \\ P_2 &= F_2(L_R - R_R), \text{ and} \\ P_3 &= F_3(L_R + R_R), \end{aligned}$$

where  $F_1$ ,  $F_2$ , and  $F_3$  are transfer functions for emphasizing the spatial content of an audio signal to achieve a perception of depth with respect to a listener upon playback of the resultant processed audio signal by a loudspeaker.

The method further comprises an act of combining the processed audio signals with the audio source signals to create the left and right output signals. The left and right output signals comprise the components recited in the following equations:

$$\begin{aligned} L_{OUT} &= K_1 L_f + K_2 L_R + K_3 P_1 + K_4 P_2 + K_5 P_3, \\ R_{OUT} &= K_6 R_f + K_7 R_R - K_8 P_1 - K_9 P_2 + K_{10} P_3, \end{aligned}$$

where  $K_1 - K_{10}$  are independent variables which determine the gain of the respective audio signal.

In another embodiment, the transfer functions  $F_1$ ,  $F_2$ , and  $F_3$  apply a level of equalization characterized by amplification of frequencies between approximately 50 and 500 Hz and between approximately 4 and 15 kHz relative to frequencies between approximately 500 Hz and 4 kHz. In yet another embodiment, the left and right output signals further comprise a center channel audio source signal. In another embodiment, the method is performed by a digital signal processing device.

According to another aspect of the invention, a method creates a simulated surround sound experience through reproduction of first and second output signals within an entertainment system having a source of at least four audio signals. The at least four audio source signals comprise a pair of front audio signals representing audio information emanating from a forward sound stage with respect to a listener, and a pair of rear audio signals representing audio information emanating from a rear sound stage with respect to the listener.

The method comprises an act of combining the front audio signals to create a front ambient component signal and a front direct component signal. The method further comprises an act of combining the rear audio signals to create a rear ambient component signal and a rear direct component signal. The method further comprises an act of processing the front ambient component signal with a first HRTF-based transfer function to create a perceived source of direction of the front ambient component about a forward left and right aspect with respect to the listener.

The method further comprises an act of processing the rear ambient component signal with a second HRTF-based transfer function to create a perceived source of direction of the rear ambient component about a rear left and right aspect with respect to the listener. The method further comprises an act of processing the rear direct component signal with a third HRTF-based transfer function to create a perceived source of direction of the rear direct component at a rear center aspect with respect to the listener.

The method further comprises an act of combining a first one of the front audio signals, a first one of the rear audio signals, the processed front ambient component, the processed rear ambient component, and the processed rear direct component to create the first output signal. The method further comprises an act of combining a second one of the front audio signals, a second one of the rear audio signals, the processed front ambient component, processed rear ambient component, and the processed rear direct component to create the second output signal. The method further comprises an act of reproducing the first and second output signals, respectively, through a pair of speakers situated in the forward sound stage with respect to the listener.

In another embodiment, the first, second, and third HRTF-based transfer functions equalize a respective inputted through amplification of signal frequencies between approximately 50 and 500 Hz and between approximately 4 and 15 kHz relative to frequencies between approximately 500 Hz and 4 kHz.

In another embodiment, the entertainment system is a personal computer system and the at least four audio source signals are generated by a digital video disk player attached to the computer system. In another embodiment, the entertainment system is a television and the at least four audio source signals are generated by an associated digital video disk player connected to the television system.

In another embodiment, the at least four audio signals comprise a center channel audio signal, the center channel signal electronically added to the first and second output signals. In another embodiment, the act of processing with the first, second, and third HRTF-based transfer functions is performed by a digital signal processor.

According to another aspect of the present invention, an audio enhancement device for use with an audio signal decoder provides multiple audio signals designated for playback through a group of speakers situated within a surround sound listening environment. The audio enhancement device generates, from the multiple audio signals, a pair of output signals for playback by a pair of speakers.

The audio enhancement device comprises an enhancement apparatus for grouping a plurality of the multiple audio signals from the signal decoder into separate pairs of audio signals. The enhancement apparatus modifies each of the separate pairs of audio signals to generate separate pairs of component signals. A circuit combines the component signals to generate enhanced audio output signals, each of the enhanced audio output signals comprising

a first component signal from a first pair of component signals and a second component signal from a second pair of component signals.

According to another aspect of the invention, an audio enhancement device for use with an audio signal decoder provides multiple audio signals designated for playback through a group of speakers situated within a surround sound listening environment. The audio enhancement device generates, from the multiple audio signals, a pair of output signals for playback by a pair of speakers.

The audio enhancement device comprises a means for grouping at least some of the multiple audio signals of the signal decoder into separate pairs of audio signals. The means for grouping, further including means for modifying each of the separate pairs of audio signals to generate separate pairs of component signals.

The audio enhancement device further comprises a means for combining the component signals to generate enhanced audio output signals. Each of the enhanced audio output signals comprise a first component signal from a first pair of component signals and a second component signal from a second pair of component signals.

#### Brief Description of the Drawings

The above and other aspects, features, and advantages of the present invention will be more apparent from the following particular description thereof presented in conjunction with the following drawings, wherein:

Figure 1 is a schematic block diagram of a first embodiment of a multi-channel audio enhancement system for generating a pair of enhanced output signals to create a surround-sound effect.

Figure 2 is a schematic block diagram of a second embodiment of a multi-channel audio enhancement system for generating a pair of enhanced output signals to create a surround-sound effect.

Figure 3 is a schematic block diagram depicting an audio enhancement process for enhancing selected pairs of audio signals.

Figure 4 is a schematic block diagram of an enhancement circuit for processing selected components from a pair of audio signals.

Figure 5 is a perspective view of a personal computer having an audio enhancement system constructed in accordance with the present invention for creating a surround-sound effect from two output signals.

Figure 6 is a schematic block diagram of the personal computer of Figure 5 depicting major internal components thereof.

Figure 7 is a diagram depicting the perceived and actual origins of sounds heard by a listener during operation of the personal computer shown in Figure 5.

Figure 8 is a schematic block diagram of a preferred embodiment for processing and mixing a group of AC-3 audio signals to achieve a surround-sound experience from a pair of output signals.

Figure 9 is a graphical representation of a first signal equalization curve for use in a preferred embodiment for processing and mixing a group of AC-3 audio signals to achieve a surround-sound experience from a pair of output signals.

11

Figure 10 is a graphical representation of a second signal equalization curve for use in a preferred embodiment for processing and mixing a group of AC-3 audio signals to achieve a surround-sound experience from a pair of output signals.

Figure 11 is a schematic block diagram depicting the various filter and amplification stages for creating the first signal equalization curve of Figure 9.

Figure 12 is a schematic block diagram depicting the various filter and amplification stages for creating the second signal equalization curve of Figure 10.

### Detailed Description of the Preferred Embodiments

Figure 1 depicts a block diagram of a first preferred embodiment of a multi-channel audio enhancement system 10 for processing a group of audio signals and providing a pair of output signals. The audio enhancement system 10 comprises a source of multi-channel audio signal source 16 which outputs a group of discrete audio signals 18 to a multi-channel signal mixer 20. The mixer 20 provides a set of processed multi-channel outputs 22 to an audio immersion processor 24. The signal processor 24 provides a processed left channel signal 26 and a processed right channel signal 28 which can be directed to a recording device 30 or to a power amplifier 32 before reproduction by a pair of speakers 34 and 36. Depending upon the signal inputs 18 received by the processor 20, the signal mixer may also generate a bass audio signal 40 containing low-frequency information which corresponds to a bass signal, B, from the signal source 16, and/or a center audio signal 42 containing dialogue or other centrally located sounds which corresponds to a center signal, C, output from the signal source 16. Not all signal sources will provide a separate bass effects channel B, nor a center channel C, and therefore it is to be understood that these channels are shown as optional signal channels. After amplification by the amplifier 32, the signals 40 and 42 are represented by the output signals 44 and 46, respectively.

In operation, the audio enhancement system 10 of Figure 1 receives audio information from the audio source 16. The audio information may be in the form of discrete analog or digital channels or as a digital data bitstream. For example, the audio source 16 may be signals generated from a group of microphones attached to various instruments in an orchestral or other audio performance. Alternatively, the audio source 16 may be a pre-recorded multi-track rendition of an audio work. In any event, the particular form of audio data received from the source 16 is not particularly relevant to the operation of the enhancement system 10.

For illustrative purposes, Figure 1 depicts the source audio signals as comprising eight main channels  $A_0$ - $A_7$ , a single bass or low-frequency channel, B, and a single center channel signal, C. It can be appreciated by one of ordinary skill in the art that the concepts of the present invention are equally applicable to any multi-channel system of greater or fewer individual audio channels.

As will be explained in more detail in connection with Figures 3 and 4, the multi-channel immersion processor 24 modifies the output signals 22 received from the mixer 20 to create an immersive three-dimensional effect when a pair of output signals,  $L_{out}$  and  $R_{out}$ , are acoustically reproduced. The processor 24 is shown in Figure 1 as an analog processor operating in real time on the multi-channel mixed output signals 22. If the processor 24

is an analog device and if the audio source 16 provides a digital data output, then the processor 24 must of course include a digital-to-analog converter (not shown) before processing the signals 22.

Referring now to Figure 2, a second preferred embodiment of a multi-channel audio enhancement system is shown which provides digital immersion processing of an audio source. An audio enhancement system 50 is shown comprising a digital audio source 52 which delivers audio information along a path 54 to a multi-channel digital audio decoder 56. The decoder 56 transmits multiple audio channel signals along a path 58. In addition, optional bass and center signals B and C may be generated by the decoder 56. Digital data signals 58, B, and C, are transmitted to an audio immersion processor 60 operating digitally to enhance the received signals. The processor 60 generates a pair of enhanced digital signals 62 and 64 which are fed to a digital to analog converter 66. In addition, the signals B and C are fed to the converter 66. The resultant enhanced analog signals 68 and 70, corresponding to the low frequency and center information, are fed to the power amplifier 32. Similarly, the enhanced analog left and right signals, 72, 74, are delivered to the amplifier 32. The left and right enhanced signals 72 and 74 may be diverted to a recording device 30 for storing the processed signals 72 and 74 directly on a recording medium such as magnetic tape or an optical disk. Once stored on recorded media, the processed audio information corresponding to signals 72 and 74 may be reproduced by a conventional stereo system without further enhancement processing to achieve the intended immersive effect described herein.

The amplifier 32 delivers an amplified left output signal 80,  $L_{OUT}$ , to the left speaker 34 and delivers an amplified right output signal 82,  $R_{OUT}$ , to the right speaker 36. Also, an amplified bass effects signal 84,  $B_{OUT}$ , is delivered to a sub-woofer 86. An amplified center signal 88,  $C_{OUT}$ , may be delivered to an optional center speaker (not shown). For near field reproductions of the signals 80 and 82, i.e., where a listener is position close to and in between the speakers 34 and 36, use of a center speaker may not be necessary to achieve adequate localization of a center image. However, in far-field applications where listeners are positioned relatively far from the speakers 34 and 36, a center speaker can be used to fix a center image between the speaker 34 and 36.

The combination consisting largely of the decoder 56 and the processor 60 is represented by the dashed line 90 which may be implemented in any number of different ways depending on a particular application, design constraints, or mere personal preference. For example, the processing performed within the region 90 may be accomplished wholly within a digital signal processor (DSP), within software loaded into a computer's memory, or as part of a micro-processor's native signal processing capabilities such as that found in Intel's Pentium generation of micro-processors.

Referring now to Figure 3, the immersion processor 24 from Figure 1 is shown in association with the signal mixer 20. The processor 24 comprises individual enhancement modules 100, 102, and 104 which each receives a pair of audio signals from the mixer 20. The enhancement modules 100, 102, and 104 process a corresponding pair of signals on the stereo level in part by isolating ambient and monophonic components from each pair of signals. These components, along with the original signals are modified to generate resultant signals 108, 110, and 112. Bass, center and other signals which undergo individual processing are delivered along a path 118 to a module 116 which may provide level adjustment, simple filtering, or other modification of the received signals

118. The resultant signals 120 from the module 116, along with the signals 108, 110, and 112 are output to a mixer 124 within the processor 24.

In Figure 4, an exemplary internal configuration of a preferred embodiment for the module 100 is depicted. The module 100 consists of inputs 130 and 132 for receiving a pair of audio signals. The audio signals are transferred to a circuit or other processing means 134 for separating the ambient components from the direct field, or monophonic, sound components found in the input signals. In a preferred embodiment, the circuit 134 generates a direct sound component along a signal path 136 representing the summation signal  $M_1 + M_2$ . A difference signal containing the ambient components of the input signals,  $M_1 - M_2$ , is transferred along a path 138. The sum signal  $M_1 + M_2$  is modified by a circuit 140 having a transfer function  $F_1$ . Similarly, the difference signal  $M_1 - M_2$  is modified by a circuit 142 having a transfer function  $F_2$ . The transfer functions  $F_1$  and  $F_2$  may be identical and in a preferred embodiment provide spatial enhancement to the inputted signals by emphasizing certain frequencies while deemphasizing others. The transfer functions  $F_1$  and  $F_2$  may also apply HRTF-based processing to the inputted signals in order to achieve a perceived placement of the signals upon playback. If desired, the circuits 140 and 142 may be used to insert time delays or phase shifts of the input signals 136 and 138 with respect to the original signals  $M_1$  and  $M_2$ .

The circuits 140 and 142 output a respective modified sum and difference signal,  $(M_1 + M_2)_p$  and  $(M_1 - M_2)_p$ , along paths 144 and 146, respectively. The original input signals  $M_1$  and  $M_2$ , as well as the processed signals  $(M_1 + M_2)_p$  and  $(M_1 - M_2)_p$  are fed to multipliers which adjust the gain of the received signals. After processing, the modified signals exit the enhancement module 100 at outputs 150, 152, 154, and 156. The output 150 delivers the signal  $K_1 M_1$ , the output 152 delivers the signal  $K_2 F_1 (M_1 + M_2)_p$ , the output 154 delivers the signal  $K_3 F_2 (M_1 - M_2)_p$ , and the output 156 delivers the signal  $K_4 M_2$ , where  $K_1 - K_4$  are constants determined by the setting of multipliers 148. The type of processing performed by the modules 100, 102, 104, and 116, and in particular the circuits 134, 140, and 142 may be user-adjustable to achieve a desired effect and/or a desired position of a reproduced sound. In some cases, it may be desirable to process only an ambient component or a monophonic component of a pair of input signals. The processing performed by each module may be distinct or it may be identical to one or more other modules.

In accordance with a preferred embodiment where a pair of audio signals is collectively enhanced before mixing, each module 100, 102, and 104 will generate four processed signals for receipt by the mixer 24 shown in Figure 3. All of the signals 108, 110, 112, and 120 may be selectively combined by the mixer 124 in accordance with principles common to one of ordinary skill in the art and dependent upon a user's preferences.

By processing multi-channel signals at the stereo level, i.e., in pairs, subtle differences and similarities within the paired signals can be adjusted to achieve an immersive effect created upon playback through speakers. This immersive effect can be positioned by applying HRTF-based transfer functions to the processed signals to create a fully immersive positional sound field. Each pair of audio signals is separately processed to create a multi-channel audio mixing system that can effectively recreate the perception of a live 360 degree sound stage. Through separate HRTF processing of the components of a pair of audio signals, e.g., the ambient and monophonic components, more

signal conditioning control is provided resulting in a more realistic immersive sound experience when the processed signals are acoustically reproduced. Examples of HRTF transfer functions which can be used to achieve a certain perceived azimuth are described in the article by E.A.B. Shaw entitled "Transformation of Sound Pressure Level From the Free Field to the Eardrum in the Horizontal Plane", J.Acoust.Soc.Am., Vol. 56, No.6, December 1974, and in the  
5 article by S. Mehrgardt and V. Mellert entitled "Transformation Characteristics of the External Human Ear", J.Acoust.Soc.Am., Vol. 61, No. 6, June 1977, both of which are incorporated herein by reference as though fully set forth.

Although principles of the present invention as described above in connection with Figures 1-4 are suitable for use in professional recording studios to make high-quality recordings, one particular application of the present  
10 invention is in audio playback devices which have the capability to process but not reproduce multi-channel audio signals. For example, today's audio-visual recorded media are being encoded with multiple audio channel signals for reproduction in a home theater surround processing system. Such surround systems typically include forward or front speakers for reproducing left and right stereo signals, rear speakers for reproducing left surround and right surround signals, a center speaker for reproducing a center signal, and a subwoofer speaker for reproduction of a low-  
15 frequency signal. Recorded media which can be played by such surround systems may be encoded with multi-channel audio signals through such techniques as Dolby's proprietary AC-3 audio encoding standard. Many of today's playback devices are not equipped with surround or center channel speakers. As a consequence, the full capability of the multi-channel recorded media may be left untapped leaving the user with an inferior listening experience.

Referring now to Figure 5, a personal computer system 200 is shown having an immersive positional audio  
20 processor constructed in accordance with the present invention. The computer system 200 consists of a processing unit 202 coupled to a display monitor 204. A front left speaker 206 and front right speaker 208, along with an optional sub-woofer speaker 210 are all connected to the unit 202 for reproducing audio signals generated by the unit 202. A listener 212 operates the computer system 200 via a keyboard 214. The computer system 200 processes a multi-channel audio signal to provide the listener 212 with an immersive 360 degree surround sound  
25 experience from just the speakers 206, 208 and the speaker 210 if available. In accordance with a preferred embodiment, the processing system disclosed herein will be described for use with Dolby AC-3 recorded media. It can be appreciated, however, that the same or similar principles may be applied to other standardized audio recording techniques which use multiple channels to create a surround sound experience. Moreover, while a computer system 200 is shown and described in Figure 5, the audio-visual playback device for reproducing the AC-3 recorded media  
30 may be a television, a combination television/personal computer, a digital video disk player coupled to a television, or any other device capable of playing a multi-channel audio recording.

Figure 6 is a schematic block diagram of the major internal components of the processing unit 202 of Figure 5. The unit 202 contains the components of a typical personal computer system, constructed in accordance with principles common to one of ordinary skill, including a central processing unit (CPU) 220, a mass storage memory  
35 and a temporary random access memory (RAM) system 222, an input/output control device 224, all interconnected via an internal bus structure. The unit 202 also contains a power supply 226 and a recorded media player/recorder



228 which may be a DVD device or other multi-channel audio source. The DVD player 228 supplies video data to a video decoder 230 for display on a monitor. Audio data from the DVD player 228 is transferred to an audio decoder 232 which supplies multiple channel digital audio data from the player 228 to an immersion processor 250. The audio information from the decoder 232 contains a left front signal, a right front signal, a left surround signal, a right surround signal, a center signal, and a low-frequency signal, all of which are transferred to the immersion audio processor 250. The processor 250 digitally enhances the audio information from the decoder 232 in a manner suitable for playback with a conventional stereo playback system. Specifically, a left channel signal 252 and a right channel signal 254 are provided as outputs from the processor 250. A low-frequency sub-woofer signal 256 is also provided for delivery of bass response in a stereo playback system. The signals 252, 254, and 256 are first provided to a digital-to-analog converter 258, then to an amplifier 260, and then output for connection to corresponding speakers.

Referring now to Figure 7, a schematic representation of speaker locations of the system of Figure 5 is shown from an overhead perspective. The listener 212 is positioned in front of and between the left front speaker 206 and the right front speaker 208. Through processing of surround signals generated from an AC-3 compatible recording in accordance with a preferred embodiment, a simulated surround experience is created for the listener 212. In particular, ordinary playback of two channel signals through the speakers 206 and 208 will create a perceived phantom center speaker 214 from which monophonic components of left and right signals will appear to emanate. Thus, the left and right signals from an AC-3 six channel recording will produce the center phantom speaker 214 when reproduced through the speakers 206 and 208. The left and right surround channels of the AC-3 six channel recording are processed so that ambient surround sounds are perceived as emanating from rear phantom speakers 215 and 216 while monophonic surround sounds appear to emanate from a rear phantom center speaker 218. Furthermore, both the left and right front signals, and the left and right surround signals, are spatially enhanced to provide an immersive sound experience to eliminate the actual speakers 206, 208 and the phantom speakers 215, 216, and 218, as perceived point sources of sound. Finally, the low-frequency information is reproduced by an optional sub-woofer speaker 210 which may be placed at any location about the listener 212.

Figure 8 is a schematic representation of an immersive processor and mixer for achieving a perceived immersive surround effect shown in Figure 7. The processor 250 corresponds to that shown in Figure 6 and receives six audio channel signals consisting of a front main left signal  $M_L$ , a front main right signal  $M_R$ , a left surround signal  $S_L$ , a right surround signal  $S_R$ , a center channel signal  $C$ , and a low-frequency effects signal  $B$ . The signals  $M_L$  and  $M_R$  are fed to corresponding gain-adjusting multipliers 252 and 254 which are controlled by a volume adjustment signal  $M_{\text{volume}}$ . The gain of the center signal  $C$  may be adjusted by a first multiplier 256, controlled by the signal  $M_{\text{volume}}$  and a second multiplier 258 controlled by a center adjustment signal  $C_{\text{volume}}$ . Similarly, the surround signals  $S_L$  and  $S_R$  are first fed to respective multipliers 260 and 262 which are controlled by a volume adjustment signal  $S_{\text{volume}}$ .

The main front left and right signals,  $M_L$  and  $M_R$ , are each fed to summing junctions 264 and 266. The summing junction 264 has an inverting input which receives  $M_R$  and a non-inverting input which receives  $M_L$  which

combine to produce  $M_L \cdot M_R$  along an output path 268. The signal  $M_L \cdot M_R$  is fed to an enhancement circuit 270 which is characterized by a transfer function  $P_1$ . A processed difference signal,  $(M_L \cdot M_R)_p$ , is delivered at an output of the circuit 270 to a gain adjusting multiplier 272. The output of the multiplier 272 is fed directly to a left mixer 280 and to an inverter 282. The inverted difference signal  $(M_R \cdot M_L)_p$  is transmitted from the inverter 282 to a right mixer 284. A summation signal  $M_L + M_R$  exits the junction 266 and is fed to a gain adjusting multiplier 286. The output of the multiplier 286 is fed to a summing junction which adds the center channel signal,  $C$ , with the signal  $M_L + M_R$ . The combined signal,  $M_L + M_R + C$ , exits the junction 290 and is directed to both the left mixer 280 and the right mixer 284. Finally, the original signals  $M_L$  and  $M_R$  are first fed through fixed gain adjustment circuits, i.e., amplifiers, 290 and 292, respectively, before transmission to the mixers 280 and 284.

The surround left and right signals,  $S_L$  and  $S_R$ , exit the multipliers 260 and 262, respectively, and are each fed to summing junctions 300 and 302. The summing junction 300 has an inverting input which receives  $S_R$  and a non-inverting input which receives  $S_L$  which combine to produce  $S_L \cdot S_R$  along an output path 304. All of the summing junctions 264, 266, 300, and 302 may be configured as either an inverting amplifier or a non-inverting amplifier, depending on whether a sum or difference signal is generated. Both inverting and non-inverting amplifiers may be constructed from ordinary operational amplifiers in accordance with principles common to one of ordinary skill in the art. The signal  $S_L \cdot S_R$  is fed to an enhancement circuit 306 which is characterized by a transfer function  $P_2$ . A processed difference signal,  $(S_L \cdot S_R)_p$ , is delivered at an output of the circuit 306 to a gain adjusting multiplier 308. The output of the multiplier 308 is fed directly to the left mixer 280 and to an inverter 310. The inverted difference signal  $(S_R \cdot S_L)_p$  is transmitted from the inverter 310 to the right mixer 284. A summation signal  $S_L + S_R$  exits the junction 302 and is fed to a separate enhancement circuit 320 which is characterized by a transfer function  $P_3$ . A processed summation signal,  $(S_L + S_R)_p$ , is delivered at an output of the circuit 320 to a gain adjusting multiplier 332. While reference is made to sum and difference signals, it should be noted that use of actual sum and difference signals is only representative. The same processing can be achieved regardless of how the ambient and monophonic components of a pair of signals are isolated. The output of the multiplier 332 is fed directly to the left mixer 280 and to the right mixer 284. Also, the original signals  $S_L$  and  $S_R$  are first fed through fixed-gain amplifiers 330 and 334, respectively, before transmission to the mixers 280 and 284. Finally, the low-frequency effects channel,  $B$ , is fed through an amplifier 336 to create the output low-frequency effects signal,  $B_{OUT}$ . Optionally, the low frequency channel,  $B$ , may be mixed as part of the output signals,  $L_{OUT}$  and  $R_{OUT}$ , if no subwoofer is available.

The enhancement circuit 250 of Figure 8 may be implemented in an analog discrete form, in a semiconductor substrate, through software run on a main or dedicated microprocessor, within a digital signal processing (DSP) chip, i.e., firmware, or in some other digital format. It is also possible to use a hybrid circuit structure combining both analog and digital components since in many cases the source signals will be digital. Accordingly, an individual amplifier, an equalizer, or other components, may be realized by software or firmware. Moreover, the enhancement circuit 270 of Figure 8, as well as the enhancement circuits 306 and 320, may employ a variety of audio enhancement techniques. For example, the circuit devices 270, 306, and 320 may use time-delay techniques, phase-shift techniques, signal equalization, or a combination of all of these techniques to achieve a

desired audio effect. The basic principles of such audio enhancement techniques are common to one of ordinary skill in the art.

In a preferred embodiment, the immersion processor circuit 250 uniquely conditions a set of AC-3 multi-channel signals to provide a surround sound experience through playback of the two output signals  $L_{OUT}$  and  $R_{OUT}$ . Specifically, the signals  $M_L$  and  $M_R$  are processed collectively by isolating the ambient information present in these signals. The ambient signal component represents the differences between a pair of audio signals. An ambient signal component derived from a pair of audio signals is therefore often referred to as the "difference" signal component. While the circuits 270, 306, and 320 are shown and described as generating sum and difference signals, other embodiments of audio enhancement circuits 270, 306, and 320 may not distinctly generate sum and difference signals at all. This can be accomplished in any number of ways using ordinary circuit design principles. For example, the isolation of the difference signal information and its subsequent equalization may be performed digitally, or performed simultaneously at the input stage of an amplifier circuit. In addition to processing of AC-3 audio signal sources, the circuit 250 of Figure 8 will automatically process signal sources having fewer discrete audio channels. For example, if Dolby Pro-Logic signals are input by the processor 250, i.e., where  $S_L = S_R$ , only the enhancement circuit 320 will operate to modify the rear channel signals since no ambient component will be generated at the junction 300. Similarly, if only two-channel stereo signals,  $M_L$  and  $M_R$ , are present, then the processor 250 operates to create a spatially enhanced listening experience from only two channels through operation of the enhancement circuit 270.

In accordance with a preferred embodiment, the ambient information of the front channel signals, which can be represented by the difference  $M_L - M_R$ , is equalized by the circuit 270 according to the frequency response curve 350 of Figure 9. The curve 350 can be referred to as a spatial correction, or "perspective", curve. Such equalization of the ambient signal information broadens and blends a perceived sound stage generated from a pair of audio signals by selectively enhancing the sound information that provides a sense of spaciousness.

The enhancement circuits 306 and 320 modify the ambient and monophonic components, respectively, of the surround signals  $S_L$  and  $S_R$ . In accordance with a preferred embodiment, the transfer functions  $P_2$  and  $P_3$  are equal and both apply the same level of perspective equalization to the corresponding input signal. In particular, the circuit 306 equalizes an ambient component of the surround signals, represented by the signal  $S_L - S_R$ , while the circuit 320 equalizes an monophonic component of the surround signals, represented by the signal  $S_L + S_R$ . The level of equalization is represented by the frequency response curve 352 of Figure 10.

The perspective equalization curves 350 and 352 are displayed in Figures 9 and 10, respectively, as a function of gain, measured in decibels, against audible frequencies displayed in log format. The gain level in decibels at individual frequencies are only relevant as they relate to a reference signal since final amplification of the overall output signals occurs in the final mixing process. Referring initially to Figure 9, and according to a preferred embodiment, the perspective curve 350 has a peak gain at a point A located at approximately 125 Hz. The gain of the perspective curve 350 decreases above and below 125 Hz at a rate of approximately 6 dB per octave. The perspective curve 350 reaches a minimum gain at a point B within a range of approximately 1.5 - 2.5 kHz. The gain

increases at frequencies above point B at a rate of approximately 6 dB per octave up to a point C at approximately 7 kHz, and then continues to increase up to approximately 20 kHz, i.e., approximately the highest frequency audible to the human ear.

Referring now to Figure 10, and according to a preferred embodiment, the perspective curve 352 has a peak  
5 gain at a point A located at approximately 125 Hz. The gain of the perspective curve 350 decreases below 125 Hz at a rate of approximately 6 dB per octave and decreases above 125 Hz at a rate of approximately 6 dB per octave. The perspective curve 352 reaches a minimum gain at a point B within a range of approximately 1.5 - 2.5 kHz. The gain increases at frequencies above point B at a rate of approximately 6 dB per octave up to a maximum-gain point C at approximately 10.5 - 11.5 kHz. The frequency response of the curve 352 decreases at frequencies  
10 above approximately 11.5 kHz.

Apparatus and methods suitable for implementing the equalization curves 350 and 352 of Figures 9 and 10 are similar to those disclosed in pending application serial number 08/430751 filed on April 27, 1995, which is incorporated herein by reference as though fully set forth. Related audio enhancement techniques for enhancing ambient information are disclosed in U.S. Patent Nos. 4,738,669 and 4,866,744, issued to Arnold I. Klayman, both  
15 of which are also incorporated by reference as though fully set forth herein.

In operation, the circuit 250 of Figure 8 uniquely functions to position the five main channel signals,  $M_L$ ,  $M_R$ ,  $C$ ,  $S_L$ , and  $S_R$ , about a listener upon reproduction by only two speakers. As discussed previously, the curve 350 of Figure 9 applied to the signal  $M_L-M_R$  broadens and spatially enhances ambient sounds from the signals  $M_L$  and  $M_R$ . This creates the perception of a wide forward sound stage emanating from the speakers 206 and 208 shown in  
20 Figure 7. This is accomplished through selective equalization of the ambient signal information to emphasize the low and high frequency components. Similarly, the equalization curve 352 of Figure 10 is applied to the signal  $S_L-S_R$  to broaden and spatially enhance the ambient sounds from the signals  $S_L$  and  $S_R$ . In addition, however, the equalization curve 352 modifies the signal  $S_L-S_R$  to account for HRTF positioning to obtain the perception of rear speakers 215 and 216 of Figure 7. As a result, the curve 352 contains a higher level of emphasis of the low and high frequency  
25 components of the signal  $S_L-S_R$  with respect to that applied to  $M_L-M_R$ . This is required since the normal frequency response of the human ear for sounds directed at a listener from zero degrees azimuth will emphasize sounds centered around approximately 2.75 kHz. The emphasis of these sounds results from the inherent transfer function of the average human pinna and from ear canal resonance. The perspective curve 352 of Figure 10 counteracts the inherent transfer function of the ear to create the perception of rear speakers for the signals  $S_L-S_R$  and  $S_L+S_R$ . The resultant processed difference signal  $(S_L-S_R)_p$  is driven out of phase to the corresponding mixers 280 and 284 to  
30 maintain the perception of a broad rear sound stage as if reproduced by phantom speakers 215 and 216.

By separating the surround signal processing into sum and difference components, greater control is provided by allowing the gain of each signal,  $S_L-S_R$  and  $S_L+S_R$ , to be adjusted separately. The present invention also recognizes that creation of a center rear phantom speaker 218, as shown in Figure 7, requires similar processing  
35 of the sum signal  $S_L+S_R$  since the sounds actually emanate from forward speakers 206 and 208. Accordingly, the signal  $S_L+S_R$  is also equalized by the circuit 320 according to the curve 352 of Figure 10. The resultant processed

signal  $(S_L + S_R)_p$  is driven in-phase to achieve the perceived phantom speaker 218 as if the two phantom rear speakers 215 and 216 actually existed. For audio reproduction systems which include a dedicated center channel speaker, the circuit 250 of Figure 8 can be modified so that the center signal C is fed directly to such center speaker instead of being mixed at the mixers 280 and 284.

5 The approximate relative gain values of the various signals within the circuit 250 can be measured against a 0dB reference for the difference signals exiting the multipliers 272 and 308. With such a reference, the gain of the amplifiers 290, 292, 330, and 334 in accordance with a preferred embodiment is approximately -18 dB, the gain of the sum signal exiting the amplifier 332 is approximately -20 dB, the gain of the sum signal exiting the amplifier 286 is approximately -20 dB, and the gain of the center channel signal exiting the amplifier 258 is approximately -

10 7 dB. These relative gain values are purely design choices based upon user preferences and may be varied without departing from the spirit of the invention. Adjustment of the multipliers 272, 286, 308, and 332 allows the processed signals to be tailored to the type of sound reproduced and tailored to a user's personal preferences. An increase in the level of a sum signal emphasizes the audio signals appearing at a center stage positioned between a pair of speakers. Conversely, an increase in the level of a difference signal emphasizes the ambient sound

15 information creating the perception of a wider sound image. In some audio arrangements where the parameters of music type and system configuration are known, or where manual adjustment is not practical, the multipliers 272, 286, 308, and 332 may be preset and fixed at desired levels. In fact, if the level adjustment of multipliers 308 and 332 are desirably with the rear signal input levels, then it is possible to connect the enhancement circuits directly to the input signals  $S_L$  and  $S_R$ . As can be appreciated by one of ordinary skill in the art, the final ratio of individual

20 signal strength for the various signals of Figure 8 is also affected by the volume adjustments and the level of mixing applied by the mixers 280 and 284.

Accordingly, the audio output signals  $L_{OUT}$  and  $R_{OUT}$  produce a much improved audio effect because ambient sounds are selectively emphasized to fully encompass a listener within a reproduced sound stage. Ignoring the relative gains of the individual components, the audio output signals  $L_{OUT}$  and  $R_{OUT}$  are represented by the following

25 mathematical formulas:

$$L_{OUT} = M_L + S_L + (M_L - M_R)_p + (S_L - S_R)_p + (M_L + M_R + C) + (S_L + S_R)_p \quad (1)$$

$$R_{OUT} = M_R + S_R + (M_R - M_L)_p + (S_R - S_L)_p + (M_L + M_R + C) + (S_L + S_R)_p \quad (2)$$

30 The enhanced output signals represented above may be magnetically or electronically stored on various recording media, such as vinyl records, compact discs, digital or analog audio tape, or computer data storage media. Enhanced audio output signals which have been stored may then be reproduced by a conventional stereo reproduction system to achieve the same level of stereo image enhancement.

Referring to Figure 11, a schematic block diagram is shown of a circuit for implementing the equalization curve 350 of Figure 9 in accordance with a preferred embodiment. The circuit 270 inputs the ambient signal  $M_L - M_R$ , corresponding to that found at path 268 of Figure 8. The signal  $M_L - M_R$  is first conditioned by a high-pass filter 360

35

having a cutoff frequency, or -3dB frequency, of approximately 50 Hz. Use of the filter 360 is designed to avoid over-amplification of the bass components present in the signal  $M_L-M_R$ .

The output of the filter 360 is split into three separate signal paths 362, 364, and 366 in order to spectrally shape the signal  $M_L-M_R$ . Specifically,  $M_L-M_R$  is transmitted along the path 362 to an amplifier 368 and then on to a summing junction 378. The signal  $M_L-M_R$  is also transmitted along the path 364 to a low-pass filter 370, then to an amplifier 372, and finally to the summing junction 378. Lastly, the signal  $M_L-M_R$  is transmitted along the path 366 to a high-pass filter 374, then to an amplifier 376, and then to the summing junction 378. Each of the separately conditioned signals  $M_L-M_R$  are combined at the summing junction 378 to create the processed difference signal  $(M_L-M_R)_p$ . In a preferred embodiment, the low-pass filter 370 has a cutoff frequency of approximately 200 Hz while the high-pass filter 374 has a cutoff frequency of approximately 7 kHz. The exact cutoff frequencies are not critical so long as the ambient components in a low and high frequency range, relative to those in a mid-frequency range of approximately 1 to 3 kHz, are amplified. The filters 360, 370, and 374 are all first order filters to reduce complexity and cost but may conceivably be higher order filters if the level of processing, represented in Figures 9 and 10, is not significantly altered. Also in accordance with a preferred embodiment, the amplifier 368 will have an approximate gain of one-half, the amplifier 372 will have a gain of approximately 1.4, and the amplifier 376 will have an approximate gain of unity.

The signals which exit the amplifiers 368, 372, and 376 make up the components of the signal  $(M_L-M_R)_p$ . The overall spectral shaping, i.e., normalization, of the ambient signal  $M_L-M_R$  occurs as the summing junction 378 combines these signals. It is the processed signal  $(M_L-M_R)_p$  which is mixed by the left mixer 280 (shown in Fig. 8) as part of the output signal  $L_{OUT}$ . Similarly, the inverted signal  $(M_R-M_L)_p$  is mixed by the right mixer 284 (shown in Fig. 8) as part of the output signal  $R_{OUT}$ .

Referring again to Figure 9, in a preferred embodiment, the gain separation between points A and B of the perspective curve 350 is ideally designed to be 9 dB, and the gain separation between points B and C should be approximately 6 dB. These figures are design constraints and the actual figures will likely vary depending on the actual value of components used for the circuit 270. If the gain of the amplifiers 368, 372, and 376 of Figure 11 are fixed, then the perspective curve 350 will remain constant. Adjustment of the amplifier 368 will tend to adjust the amplitude level of point B thus varying the gain separation between points A and B, and points B and C. In a surround sound environment, a gain separation much larger than 9 dB may tend to reduce a listener's perception of mid-range definition.

Implementation of the perspective curve by a digital signal processor will, in most cases, more accurately reflect the design constraints discussed above. For an analog implementation, it is acceptable if the frequencies corresponding to points A, B, and C, and the constraints on gain separation, vary by plus or minus 20 percent. Such a deviation from the ideal specifications will still produce the desired enhancement effect, although with less than optimum results.

Referring now to Figure 12, a schematic block diagram is shown of a circuit for implementing the equalization curve 352 of Figure 10 in accordance with a preferred embodiment. Although the same curve 352 is

used to shape the signals  $S_L \cdot S_R$  and  $S_L + S_R$ , for ease of discussion purposes, reference is made in Figure 12 only to the circuit enhancement device 306. In a preferred embodiment, the characteristics of the device 306 is identical to that of 320. The circuit 306 inputs the ambient signal  $S_L \cdot S_R$ , corresponding to that found at path 304 of Figure 8. The signal  $S_L \cdot S_R$  is first conditioned by a high-pass filter 380 having a cutoff frequency of approximately 50 Hz.

5 As in the circuit 270 of Figure 11, the output of the filter 380 is split into three separate signal paths 382, 384, and 386 in order to spectrally shape the signal  $S_L \cdot S_R$ . Specifically, the signal  $S_L \cdot S_R$  is transmitted along the path 382 to an amplifier 388 and then on to a summing junction 396. The signal  $S_L \cdot S_R$  is also transmitted along the path 384 to a high-pass filter 390 and then to a low-pass filter 392. The output of the filter 392 is transmitted to an amplifier 394, and finally to the summing junction 396. Lastly, the signal  $S_L \cdot S_R$  is transmitted along the path 386

10 to a low-pass filter 398, then to an amplifier 400, and then to the summing junction 396. Each of the separately conditioned signals  $S_L \cdot S_R$  are combined at the summing junction 396 to create the processed difference signal  $(S_L \cdot S_R)_p$ . In a preferred embodiment, the high-pass filter 370 has a cutoff frequency of approximately 21 kHz while the low-pass filter 392 has a cutoff frequency of approximately 8 kHz. The filter 392 serves to create the maximum-gain point C of Figure 10 and may be removed if desired. Additionally, the low-pass filter 398 has a cutoff frequency

15 of approximately 225 Hz. As can be appreciated by one of ordinary skill in the art, there are many additional filter combinations which can achieve the frequency response curve 352 shown in Figure 10 without departing from the spirit of the invention. For example, the exact number of filters and the cutoff frequencies are not critical so long as the signal  $S_L \cdot S_R$  is equalized in accordance with Figure 10. In a preferred embodiment, all of the filters 380, 390, 392, and 398 are first order filters. Also in accordance with a preferred embodiment, the amplifier 388 will have

20 an approximate gain of 0.1, the amplifier 394 will have a gain of approximately 1.8, and the amplifier 400 will have an approximate gain of 0.8. It is the processed signal  $(S_L \cdot S_R)_p$  which is mixed by the left mixer 280 (shown in Fig. 8) as part of the output signal  $L_{OUT}$ . Similarly, the inverted signal  $(S_R \cdot S_L)_p$  is mixed by the right mixer 284 (shown in Fig. 8) as part of the output signal  $R_{OUT}$ .

Referring again to Figure 10, in a preferred embodiment, the gain separation between points A and B of the perspective curve 352 is ideally designed to be 18 dB, and the gain separation between points B and C should be approximately 10 dB. These figures are design constraints and the actual figures will likely vary depending on the actual value of components used for the circuits 306 and 320. If the gain of the amplifiers 388, 394, and 400 of Figure 12 are fixed, then the perspective curve 352 will remain constant. Adjustment of the amplifier 388 will tend to adjust the amplitude level of point B of the curve 352, thus varying the gain separation between points A

25 and B, and points B and C.

Through the foregoing description and accompanying drawings, the present invention has been shown to have important advantages over current audio reproduction and enhancement systems. While the above detailed description has shown, described, and pointed out the fundamental novel features of the invention, it will be understood that various omissions and substitutions and changes in the form and details of the device illustrated may

30 be made by those skilled in the art, without departing from the spirit of the invention. Therefore, the invention should be limited in its scope only by the following claims.

**WHAT IS CLAIMED IS:**

1. A system for processing at least four discrete audio signals including main left and right signals containing audio information intended for playback from a front sound stage, and surround left and right signals containing audio information intended for playback from a rear sound stage, said system generating a pair of left  
5 and right output signals for reproduction from the front sound stage to create the perception of a three dimensional sound image without the need for actual speakers placed in the rear sound stage, said system comprising:

a first electronic audio enhancer receiving said main left and right signals, said first audio enhancer processing an ambient component of said main left and right signals to create the perception of a broadened sound image across the front sound stage when said left and right output signals are reproduced by a pair  
10 of speakers positioned within the front sound stage;

a second electronic audio enhancer receiving said surround left and right signals, said second audio enhancer processing an ambient component of said surround left and right signals to create the perception of an acoustic sound image across the rear sound stage when said left and right output signals are reproduced by the pair of speakers positioned within the front sound stage;

a third electronic audio enhancer receiving said surround left and right signals, said third audio enhancer processing a monophonic component of said surround left and right signals to create the perception of an acoustic sound image at a center location of the rear sound stage when said left and right output signals are reproduced by the pair of speakers positioned within the front sound stage; and

a signal mixer for generating said left and right output signals from the at least four discrete audio signals by combining the processed ambient component from the main left and right signals, the processed ambient component for the surround left and right signals, and the processed monophonic component from the surround left and right signals, wherein said ambient components of said main and surround signals are included in the left and right output signals in an out-of-phase relationship with respect to each other.

2. The system of Claim 1 wherein said at least four discrete audio signals comprise a center channel  
25 signal containing audio information intended for playback by a front sound stage center speaker, and wherein said center channel signal is combined by said signal mixer as part of said left and right output signals.

3. The system of Claim 1 wherein said at least four discrete audio signals comprise a center channel signal containing audio information intended for playback by a center speaker located within the front sound stage, and wherein said center channel signal is combined with a monophonic component of the main left and right signals  
30 by said signal mixer to generate said left and right output signals.

4. The system of Claim 1 wherein said at least four discrete audio signals comprises a center channel signal having center stage audio information which is acoustically reproduced by a dedicated center channel speaker.

5. The system of Claim 1 wherein said first, second, and third electronic audio enhancers apply an HRTF-based transfer function to a respective one of said discrete audio signals for creating an apparent sound image  
35 corresponding to said discrete audio signals when said left and right output signals are acoustically reproduced.



23

6. The system of Claim 1 wherein said first audio enhancer equalizes said ambient component of said main left and right signals by boosting said ambient component below approximately 1 kHz and above approximately 2 kHz relative to frequencies between approximately 1 and 2 kHz.

7. The system of Claim 6 wherein the peak gain applied to boost said ambient component, relative to the gain applied to said ambient component between approximately 1 and 2 kHz, is approximately 8 dB.

8. The system of Claim 1 wherein said second and third audio enhancers equalize said ambient and monophonic components of said surround left and right signals by boosting said ambient and monophonic components below approximately 1 kHz and above approximately 2 kHz, relative to frequencies between approximately 1 and 2 kHz.

9. The system of Claim 8 wherein the peak gain applied to boost said ambient and monophonic components of said surround left and right signals, relative to the gain applied to said ambient and monophonic components between approximately 1 and 2 kHz, is approximately 18 dB.

10. The system of Claim 1 wherein said first, second, and third electronic audio enhancers are formed upon a semiconductor substrate.

11. The system of Claim 1 wherein said first, second, and third electronic audio enhancers are implemented in software.

12. A multi-channel recording and playback apparatus receives a plurality of individual audio signals and processes said plurality of audio signals to provide first and second enhanced audio output signals for achieving an immersive sound experience upon playback of said output signals, said multi-channel recording apparatus comprising:

a plurality of parallel audio signal processing devices for modifying the signal content of said individual audio signals wherein each parallel audio signal processing device comprises:

a circuit for receiving two of said individual audio signals and isolating an ambient component of said two audio signals from a monophonic component of said two audio signals;

positional processing means capable of electronically applying a head related transfer function to each of said ambient and monophonic components of said two audio signals to generate processed ambient and monophonic components, said head related transfer functions corresponding to a desired spatial location with respect to a listener; and

a multi-channel circuit mixer for combining said processed monophonic components and ambient components generated by said plurality of positional processing means to generate said enhanced audio output signals wherein said processed ambient components are combined in an out-of-phase relationship with respect to said first and second output signals.

13. The multi-channel recording and playback apparatus of Claim 12 wherein each of said plurality of positional processing means further includes a circuit capable of individually modifying said two audio signals and wherein said multi-channel mixer further combines said two modified signals from said plurality of positional processing means with said respective ambient and monophonic components to generate said audio output signals.

14. The multi-channel recording and playback apparatus of Claim 13 wherein said circuit capable of individually modifying said two audio signals electronically applies a head related transfer function to said two audio signals.

5 15. The multi-channel recording and playback apparatus of Claim 13 wherein said circuit capable of individually modifying said two audio signals electronically applies a time delay to one of said two audio signals.

16. The multi-channel recording and playback apparatus of Claim 12 wherein said two audio signals comprise audio information corresponding to a left front location and a right front location with respect to a listener.

17. The multi-channel recording and playback apparatus of Claim 12 wherein said two audio signals comprise audio information corresponding to a left rear location and a right rear location with respect to a listener.

10 18. The multi-channel recording and playback apparatus of Claim 12 wherein said plurality of parallel processing devices comprises first and second processing devices, said first processing device applying a head related transfer function to a first pair of said audio signals for achieving a first perceived direction for said first pair of audio signals when said output signals are reproduced, and said second processing device applying a head related transfer function to a second pair of said audio signals for achieving a second perceived direction for said second pair of audio signals when said output signals are reproduced.

15 19. The multi-channel recording and playback apparatus of Claim 12 wherein said plurality of parallel audio processing devices and said multi-channel circuit mixer are implemented in a digital signal processing device of said multi-channel recording and playback apparatus.

20 20. An audio enhancement system for processing a plurality of audio source signals to create a pair of stereo output signals for generating a three dimensional sound field when said pair of stereo output signals are reproduced by a pair of loudspeakers, said audio enhancement system comprising:

25 a first processing circuit in communication with a first pair of said audio source signals, said first processing circuit configured to isolate a first ambient component and a first monophonic component from said first pair of audio signals, said first processing circuit further configured to modify said first ambient component and said first monophonic component to create a first acoustic image such that said first acoustic image is perceived by a listener as emanating from a first location;

30 a second processing circuit in communication with a second pair of said audio source signals, said second processing circuit configured to isolate a second ambient component and a second monophonic component from said second pair of audio signals, said second processing circuit further configured to modify said second ambient component and said second monophonic component to create a second acoustic image, such that said second acoustic image is perceived by said listener as emanating from a second location; and

35 a mixing circuit in communication with said first processing circuit and said second processing circuit, said mixing circuit configured to combine said first and second modified monophonic components in phase and combine said first and second modified ambient components out of phase to generate a pair of stereo output signals.

21. The system of Claim 20 wherein said first processing circuit is further configured to modify a plurality of frequency components in said first ambient component with a first transfer function.

22. The system of Claim 21 wherein said first transfer function is further configured to emphasize a portion of the low frequency components in said first ambient component relative to other frequency components  
5 in said first ambient component.

23. The system of Claim 21 wherein said first transfer function is configured to emphasize a portion of the high frequency components of said first ambient component relative to other frequency components in said first ambient component.

24. The system of Claim 21 wherein said second processing circuit is configured to modify a plurality  
10 of frequency components in said second ambient component with a second transfer function.

25. The system of Claim 24 wherein said second transfer function is configured to modify said frequency components in said second ambient component in a different manner than said first transfer function modifies said frequency components in said first ambient component.

26. The system of Claim 24 wherein said second transfer function is configured to deemphasize a  
15 portion of said frequency components above approximately 11.5 kHz relative to other frequency components in said second ambient component.

27. The system of Claim 24 wherein said second transfer function is configured to deemphasize a portion of said frequency components between approximately 125 Hz and approximately 2.5 khz relative to other frequency components in said second ambient component.

28. The system of Claim 24 wherein said second transfer function is configured to increase a portion  
20 of said frequency components between approximately 2.5 khz and approximately 11.5 khz relative to other frequency components in said second ambient component.

29. A multi-track audio processor receiving a plurality of separate audio signals as part of a composite audio source, said plurality of audio signals comprising at least two distinct audio signal pairs containing audio  
25 information which is desirably interpreted by a listener as emanating from distinct locations within a sound listening environment, said multi-track audio processor comprising:

first electronic means receiving a first pair of said audio signals, said first electronic means separately applying a head related transfer function to an ambient component of said first pair of audio signals for creating a first acoustic image wherein said first acoustic image is perceived by a listener as  
30 emanating from a first location;

second electronic means receiving a second pair of said audio signals, said second electronic means separately applying a head related transfer function to an ambient component and a monophonic component of said second pair of audio signals for creating a second acoustic image wherein said second acoustic image is perceived by the listener as emanating from a second location; and

means for mixing said components of said first and second pair of audio signals received from said first and second electronic means, said means for mixing combining said ambient components out of phase to generate said pair of stereo output signals.

30. An entertainment system having two main audio reproduction channels for reproducing an audio-visual recording to a user wherein said audio-visual recording comprises five discrete audio signals including a front-left signal,  $F_L$ , a front-right signal,  $F_R$ , a rear-left signal,  $R_L$ , a rear-right signal,  $R_R$ , and a center signal,  $C$ , and wherein said entertainment system achieves a surround sound experience for said user from said two main audio channels, said entertainment system comprising:

an audio-visual playback device for extracting said five discrete audio signals from said audio-visual recording;

an audio processing device for receiving said five discrete audio signals and generating said two main audio reproduction channels, said audio processing device comprising:

a first processor for equalizing an ambient component of said front signals,  $F_L$  and  $F_R$ , to obtain a spatially-corrected ambient component  $(F_L - F_R)_p$ ;

a second processor for equalizing an ambient component of said rear signals,  $R_L$  and  $R_R$ , to obtain a spatially-corrected ambient component  $(R_L - R_R)_p$ ;

a third processor for equalizing a direct-field component of said rear signals,  $R_L$  and  $R_R$ , to obtain a spatially-corrected direct-field component  $(R_L + R_R)_p$ ;

a left mixer for generating a left output signal, said left mixer combining the spatially-corrected ambient component,  $(F_L - F_R)_p$ , with said spatially-corrected ambient component,  $(R_L - R_R)_p$ , and said spatially-corrected direct-field component,  $(R_L + R_R)_p$ , to create said left output signal; and

a right mixer for generating a right output signal, said right mixer combining an inverted spatially-corrected ambient component,  $(F_R - F_L)_p$ , with an inverted spatially-corrected ambient component,  $(R_R - R_L)_p$ , and said spatially-corrected direct-field component,  $(R_L + R_R)_p$ , to create said right output signal; and

means for reproducing said left and right output signals through said two main channels in connection with playback of said audio-visual recording to create a surround sound experience for said user.

31. The entertainment system of Claim 30 wherein said center signal is input by said left mixer and combined as part of said left output signal and said center signal is combined by said right mixer and combined as part of said right output signal.

32. The entertainment system of Claim 30 wherein said center signal and a direct field component of said front signals,  $F_L + F_R$ , are combined by said left and right mixers as part of said left and right output signals, respectively.

33. The entertainment system of Claim 30 wherein said center signal is provided as a third output signal for reproduction by a center channel speaker of said entertainment system.

34. The entertainment system of Claim 30 wherein said entertainment system is a personal computer and said audio-visual playback device is a digital versatile disk (DVD) player.

35. The entertainment system of Claim 30 wherein said entertainment system is a television and said audio-visual playback device is an associated digital versatile disk (DVD) player connected to said television system.

5 36. The entertainment system of Claim 30 wherein said first, second, and third processors emphasize a low and high range of frequencies relative to a mid-range of frequencies.

37. The entertainment system of Claim 30 wherein said audio processing device is implemented as an analog circuit formed upon a semiconductor substrate.

10 38. The entertainment system of Claim 30 wherein said audio processing device is implemented in a software format, said software format executed by a microprocessor of said entertainment system.

39. A method of enhancing a group of audio source signals wherein the audio source signals are designated for speakers placed around a listener to create left and right output signals for acoustic reproduction by a pair of speakers in order to simulate a surround sound environment, the audio source signals comprising a left-front signal ( $L_F$ ), a right-front signal ( $R_F$ ), a left-rear signal ( $L_R$ ), and a right-rear signal ( $R_R$ ), said method of enhancing  
15 comprising the following steps:

modifying said audio source signals to create processed audio signals based on the audio content of selected pairs of said source signals, said processed audio signals defined in accordance with the following equations:

$$\begin{aligned} P_1 &= F_1(L_F - R_F), \\ 20 \quad P_2 &= F_2(L_R - R_R), \text{ and} \\ P_3 &= F_3(L_R + R_R), \end{aligned}$$

where  $F_1$ ,  $F_2$ , and  $F_3$  are transfer functions for emphasizing the spatial content of an audio signal to achieve a perception of depth with respect to a listener upon playback of the resultant processed audio signal by a loudspeaker, and

25 combining said processed audio signals with said audio source signals to create said left and right output signals, said left and right output signals comprising the components recited in the following equations:

$$\begin{aligned} L_{OUT} &= K_1 L_F + K_2 L_R + K_3 P_1 + K_4 P_2 + K_5 P_3, \\ 30 \quad R_{OUT} &= K_6 R_F + K_7 R_R + K_8 P_1 + K_9 P_2 + K_{10} P_3, \end{aligned}$$

where  $K_1 - K_{10}$  are independent variables which determine the gain of the respective audio signal.

40. The method of enhancing a group of audio source signals as recited in Claim 39 wherein the transfer functions  $F_1$ ,  $F_2$ , and  $F_3$  apply a level of equalization characterized by amplification of frequencies between approximately 50 and 500 Hz and between approximately 4 and 15 kHz relative to frequencies between approximately 500 Hz and 4 kHz.

35 41. The method of enhancing a group of audio source signals as recited in Claim 39 wherein the left and right output signals further comprise a center channel audio source signal.

42. The method of enhancing a group of audio source signals as recited in Claim 39 wherein said method is performed by a digital signal processing device.

43. A method of creating a simulated surround sound experience through reproduction of first and second output signals within an entertainment system having a source of at least four audio signals wherein said  
5 at least four audio source signals comprise a pair of front audio signals representing audio information emanating from a forward sound stage with respect to a listener, and a pair of rear audio signals representing audio information emanating from a rear sound stage with respect to the listener, said method comprising the following steps:

combining said front audio signals to create a front ambient component signal and a front direct component signal,

10 combining said rear audio signals to create a rear ambient component signal and a rear direct component signal,

processing the front ambient component signal with a first HRTF-based transfer function to create a perceived source of direction of said front ambient component about a forward left and right aspect with respect to the listener,

15 processing the rear ambient component signal with a second HRTF-based transfer function to create a perceived source of direction of said rear ambient component about a rear left and right aspect with respect to the listener,

processing the rear direct component signal with a third HRTF-based transfer function to create a perceived source of direction of said rear direct component at a rear center aspect with respect to the  
20 listener, and

combining a first one of said front audio signals, a first one of said rear audio signals, said processed front ambient component, said processed rear ambient component, and said processed rear direct component to create said first output signal,

25 combining a second one of said front audio signals, a second one of said rear audio signals, said processed front ambient component, said processed rear ambient component, and said processed rear direct component to create said second output signal, and

reproducing said first and second output signals, respectively, through a pair of speakers situated in said forward sound stage with respect to the listener.

44. The method of Claim 43 wherein said first, second, and third HRTF-based transfer functions  
30 equalize a respective inputted through amplification of signal frequencies between approximately 50 and 500 Hz and between approximately 4 and 15 kHz relative to frequencies between approximately 500 Hz and 4 kHz.

45. The method of Claim 43 wherein the entertainment system is a personal computer system and said at least four audio source signals are generated by a digital video disk player attached to said computer system.

46. The method of Claim 43 wherein the entertainment system is a television and said at least four  
35 audio source signals are generated by an associated digital video disk player connected to said television system.

29

47. The method of Claim 43 wherein said at least four audio signals comprise a center channel audio signal, said center channel signal electronically added to said first and second output signals.

48. The method of Claim 43 wherein said steps of processing with said first, second, and third HRTF-based transfer functions is performed by a digital signal processor.

5 49. An audio enhancement device for use with an audio signal decoder providing multiple audio signals designated for playback through a group of speakers situated within a surround sound listening environment, said audio enhancement device generating, from said multiple audio signals, a pair of output signals for playback by a pair of speakers, said device comprising:

10 an enhancement apparatus for grouping a plurality of the multiple audio signals from the signal decoder into separate pairs of audio signals, said enhancement apparatus modifying each of said separate pairs of audio signals to generate separate pairs of component signals; and

a circuit for combining said component signals to generate enhanced audio output signals, each of said enhanced audio output signals comprising a first component signal from a first pair of component signals and a second component signal from a second pair of component signals.

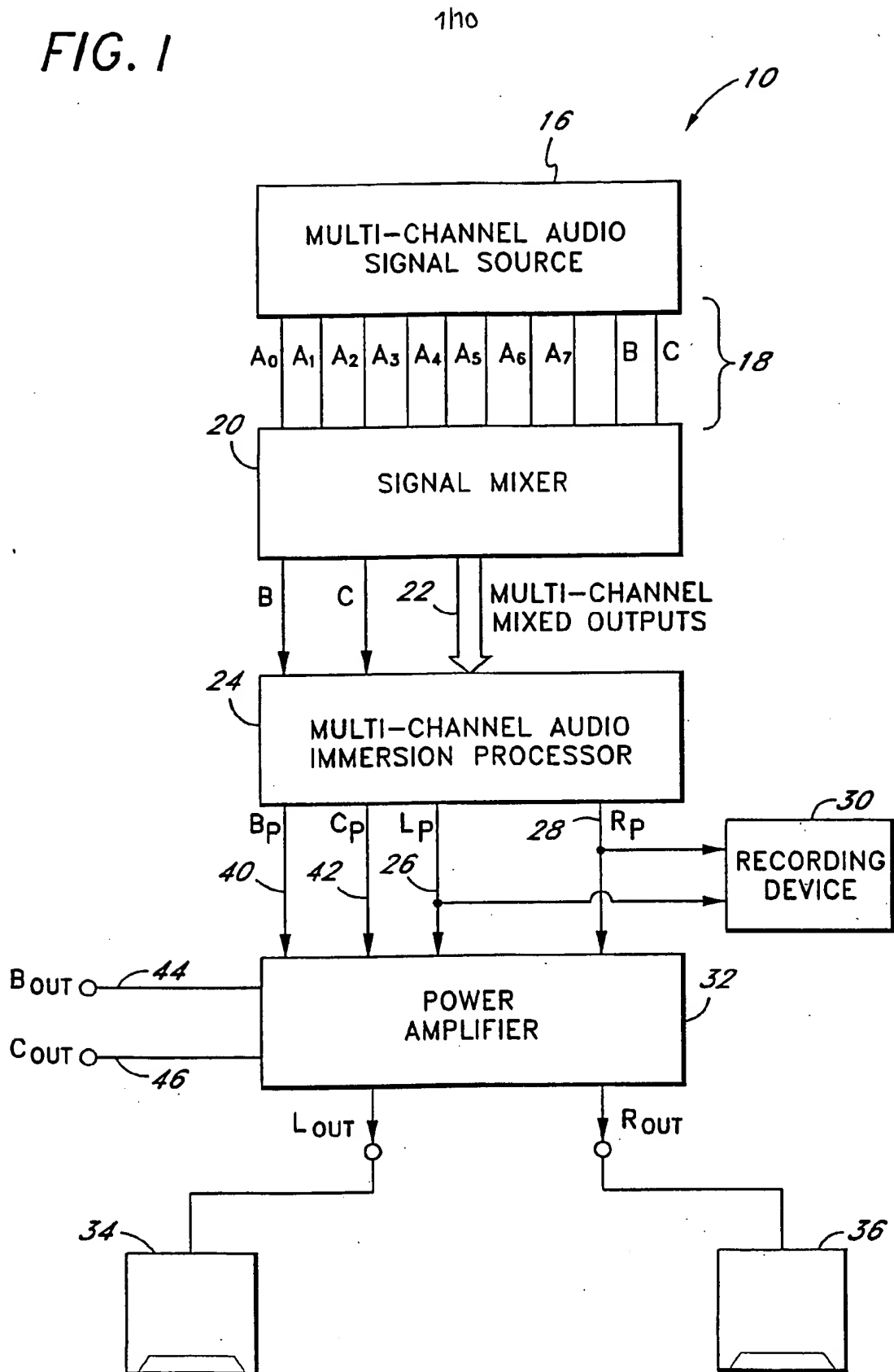
15 50. An audio enhancement device for use with an audio signal decoder providing multiple audio signals designated for playback through a group of speakers situated within a surround sound listening environment, said audio enhancement device generating, from said multiple audio signals, a pair of output signals for playback by a pair of speakers, said device comprising:

20 means for grouping at least some of the multiple audio signals of the signal decoder into separate pairs of audio signals, said means for grouping further including means for modifying each of said separate pairs of audio signals to generate separate pairs of component signals; and

means for combining said component signals to generate enhanced audio output signals, each of said enhanced audio output signals comprising a first component signal from a first pair of component signals and a second component signal from a second pair of component signals.

25

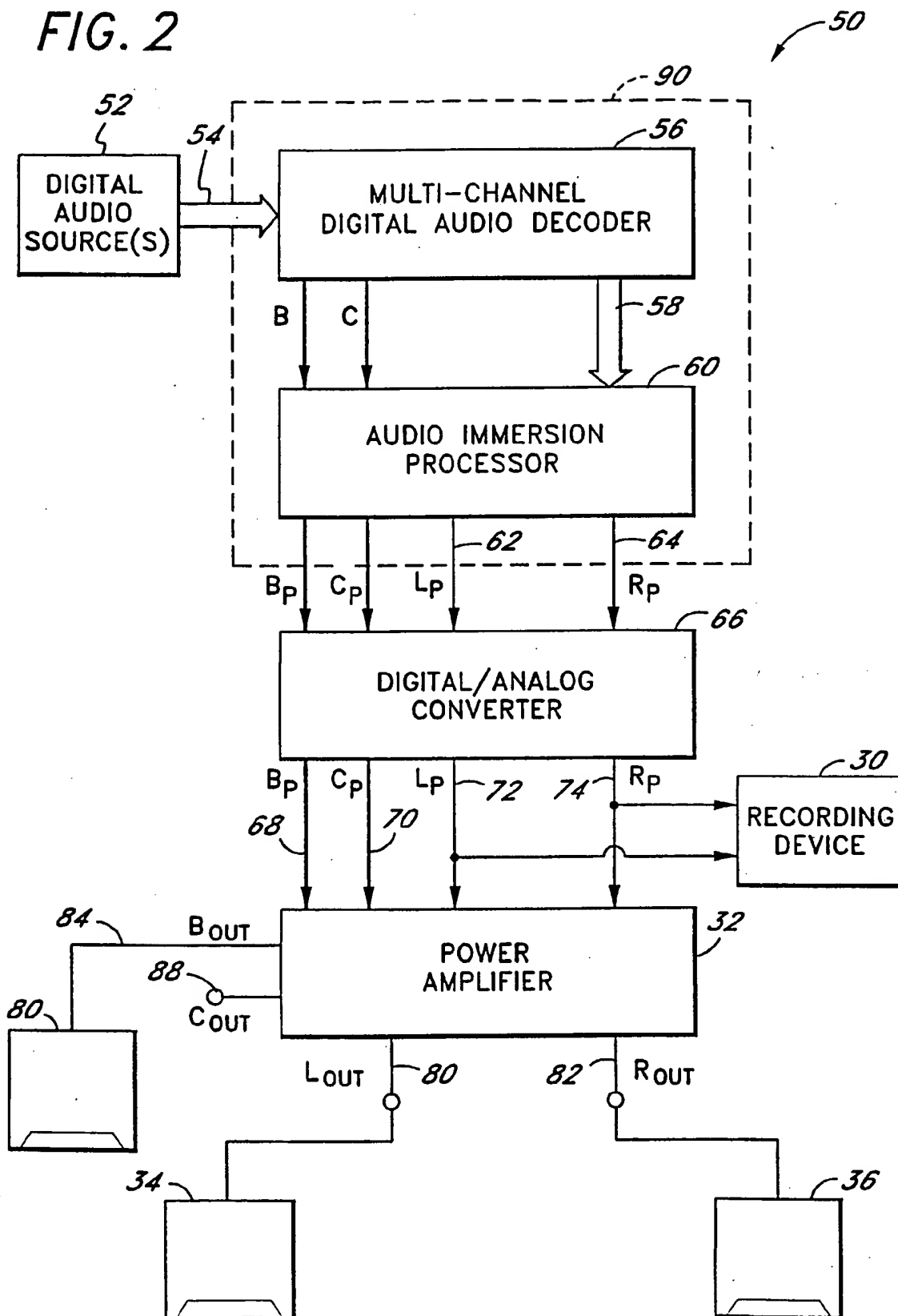
FIG. 1





2/10

FIG. 2



3/10

FIG. 3

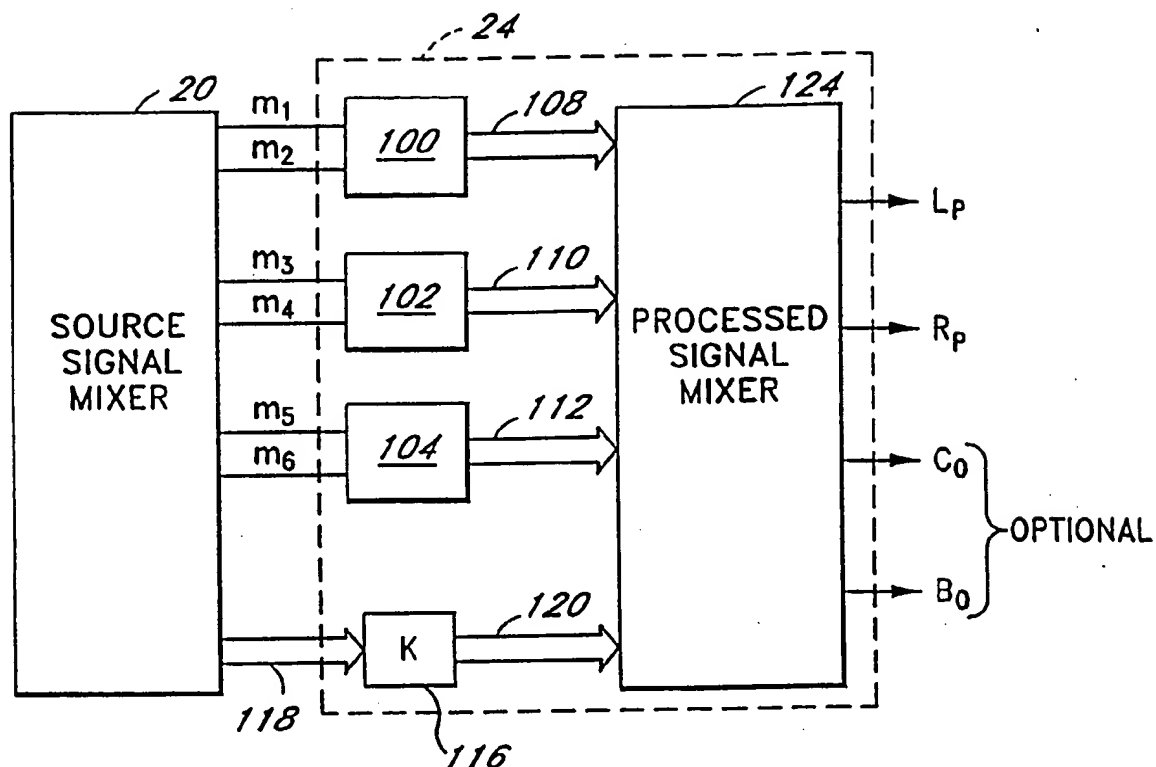
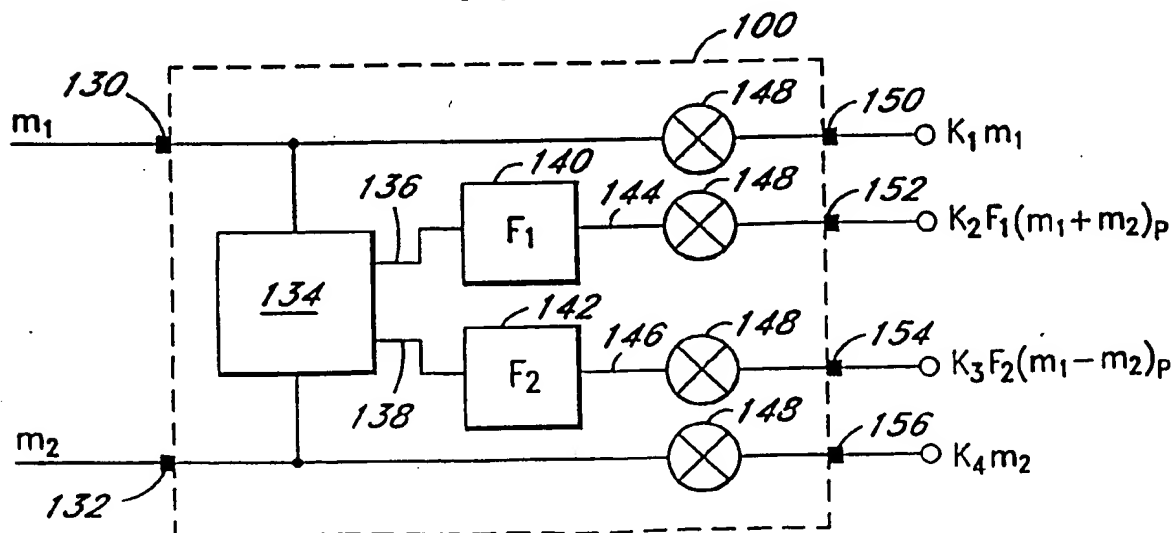


FIG. 4



4/10

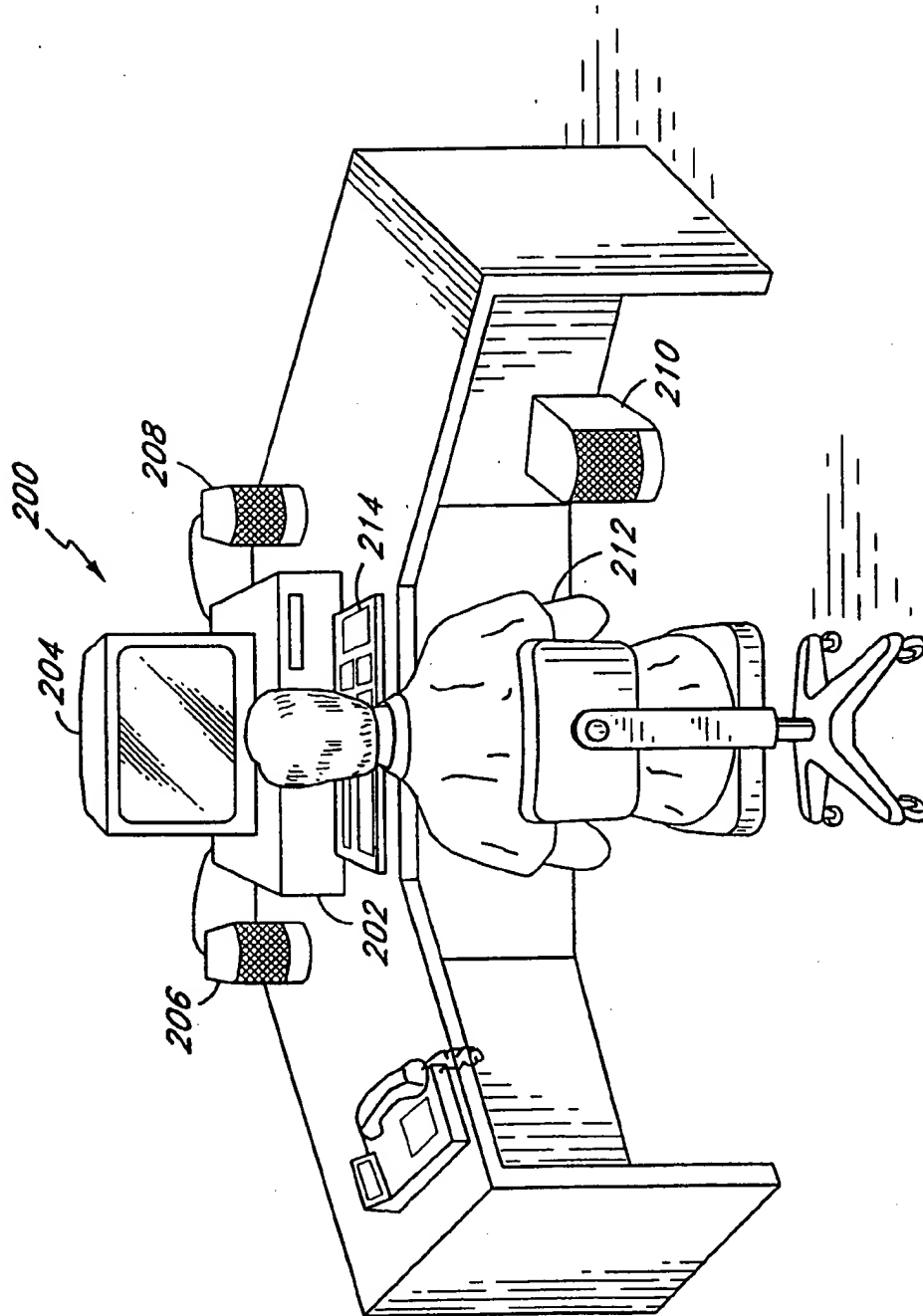


FIG. 5

5/10

FIG. 6

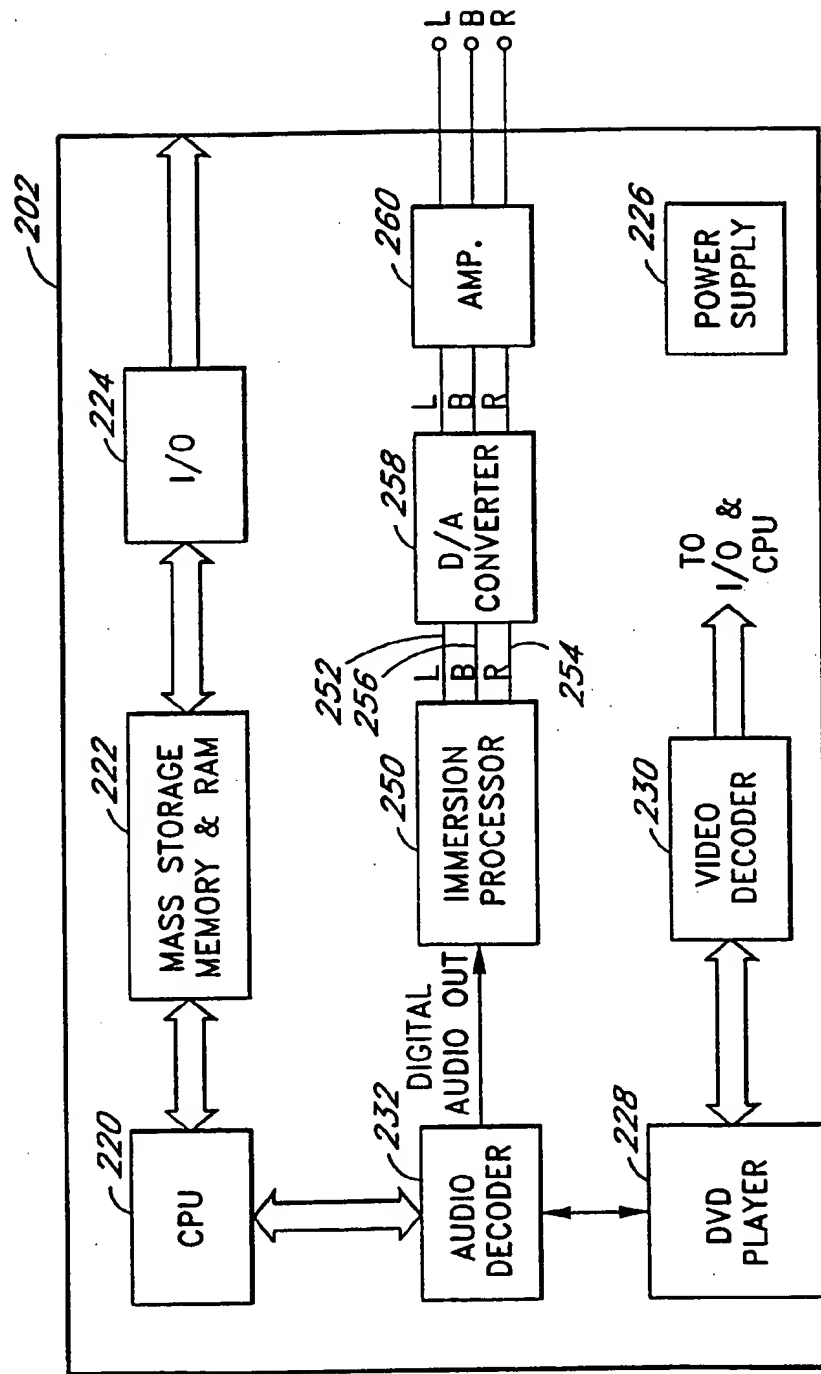
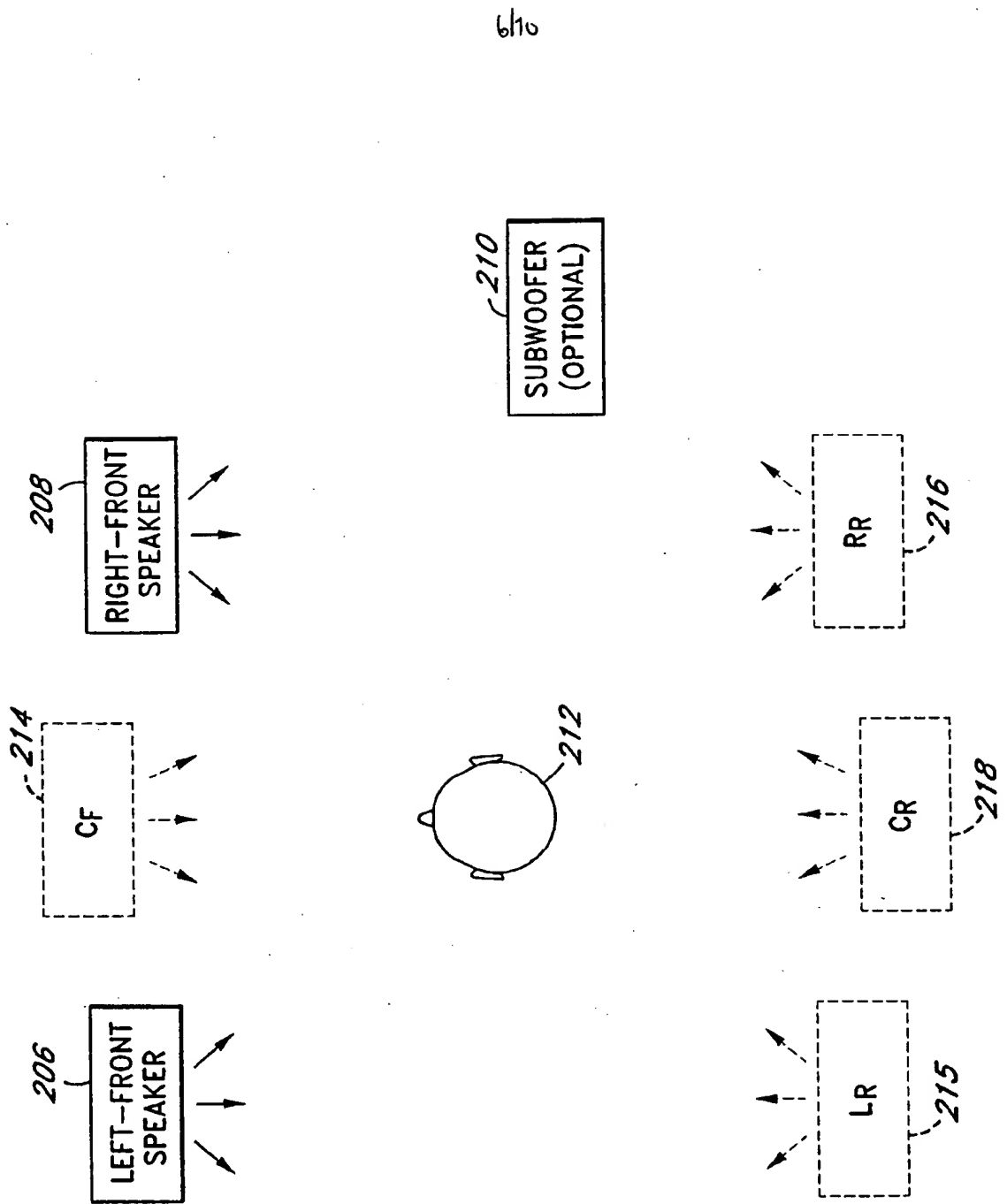
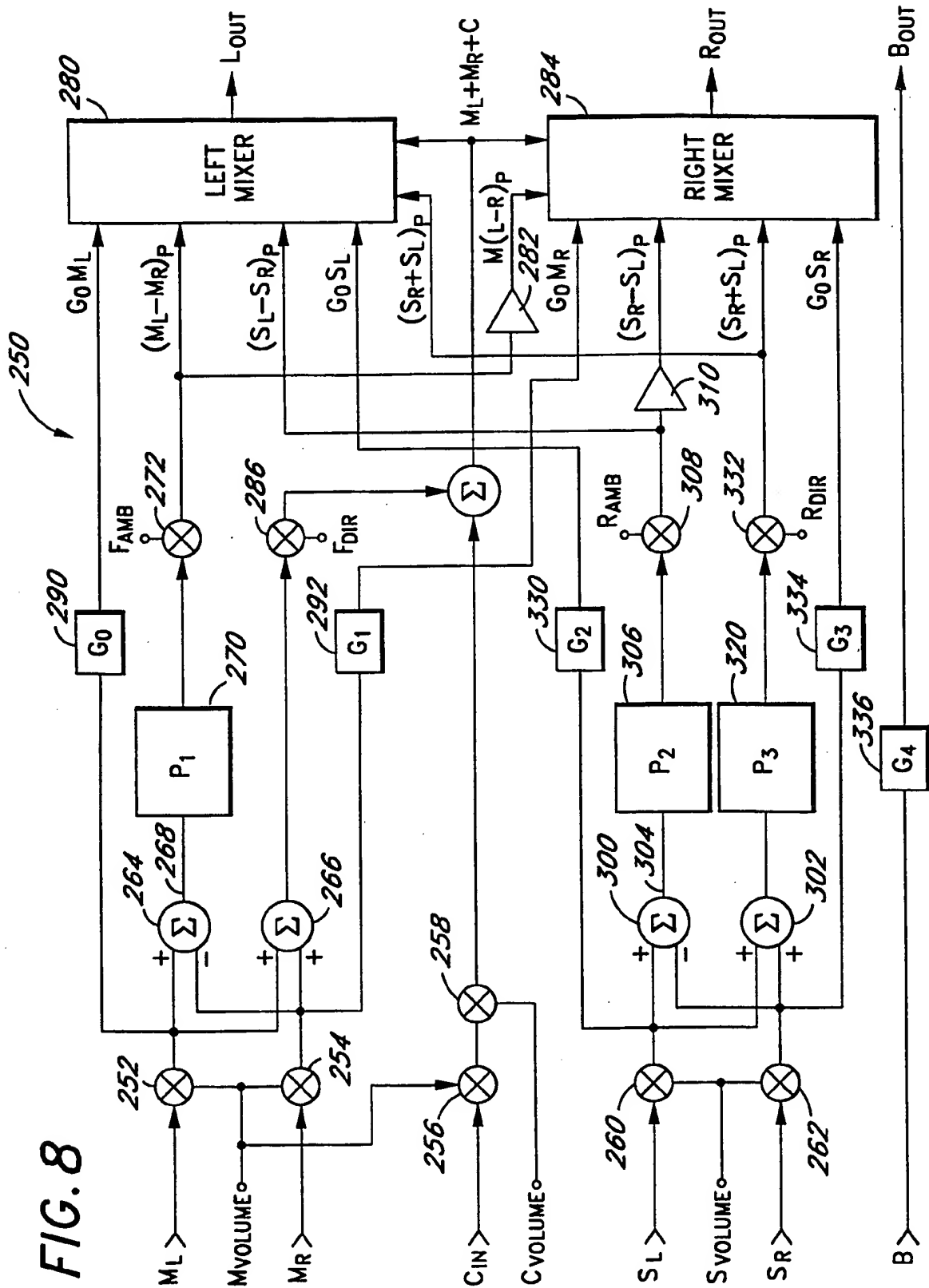


FIG. 7

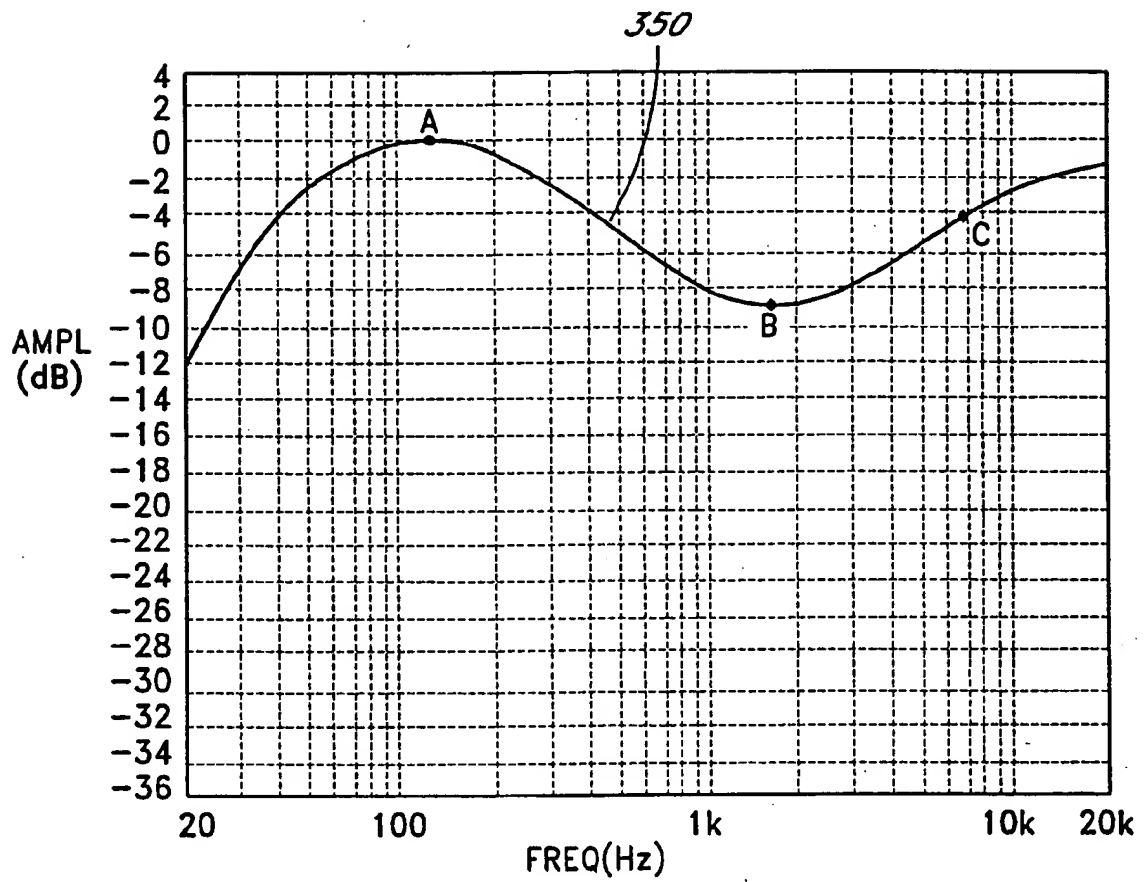


7/10



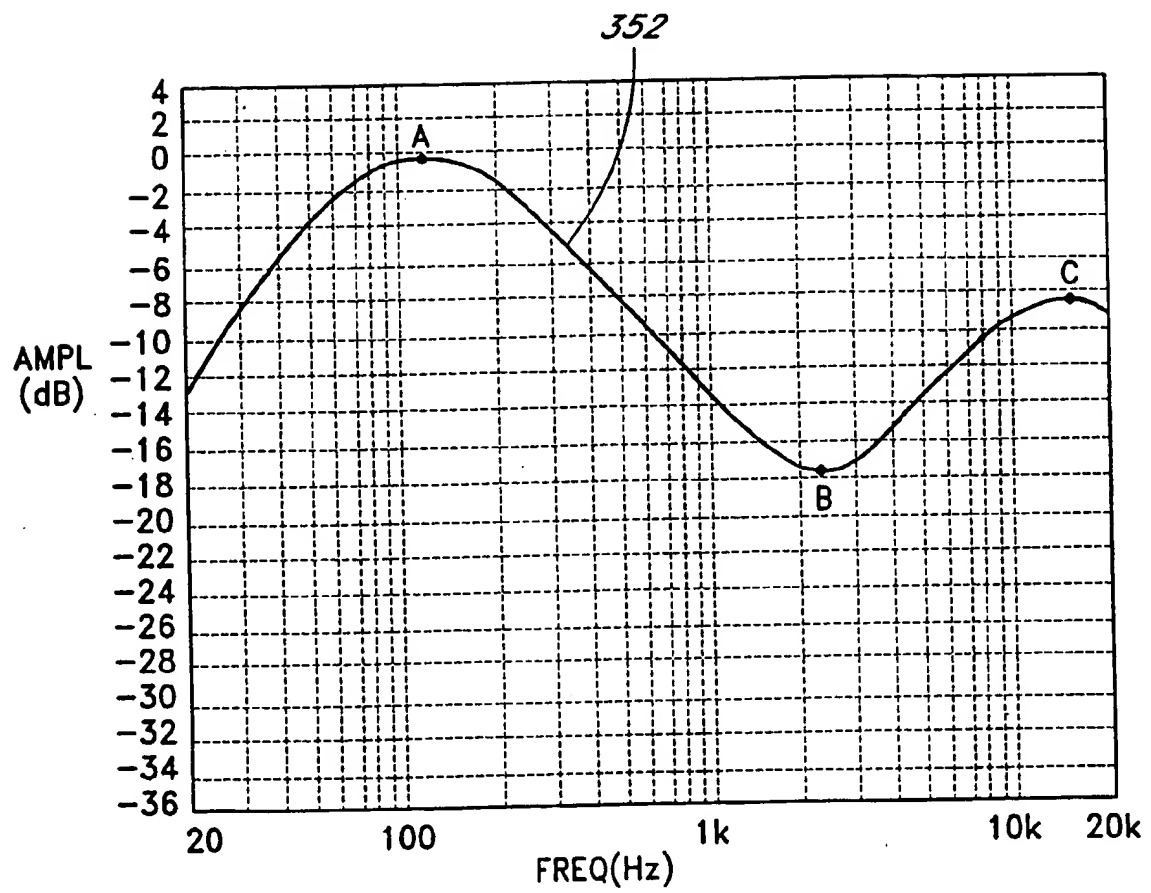
810

FIG. 9



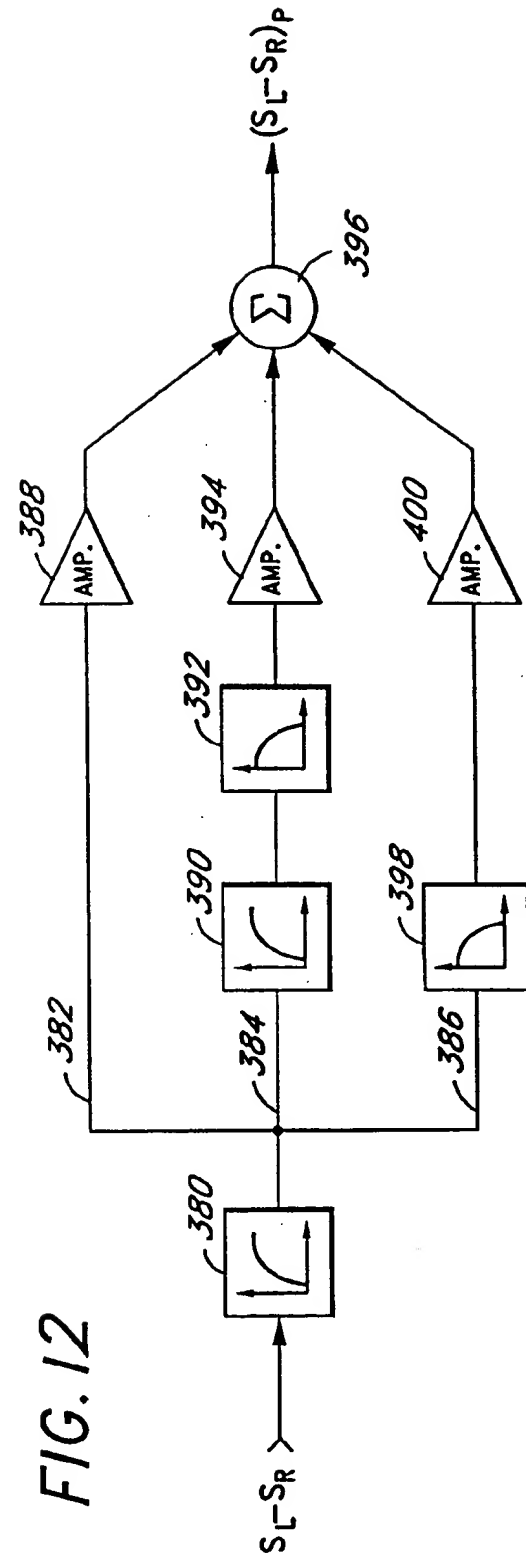
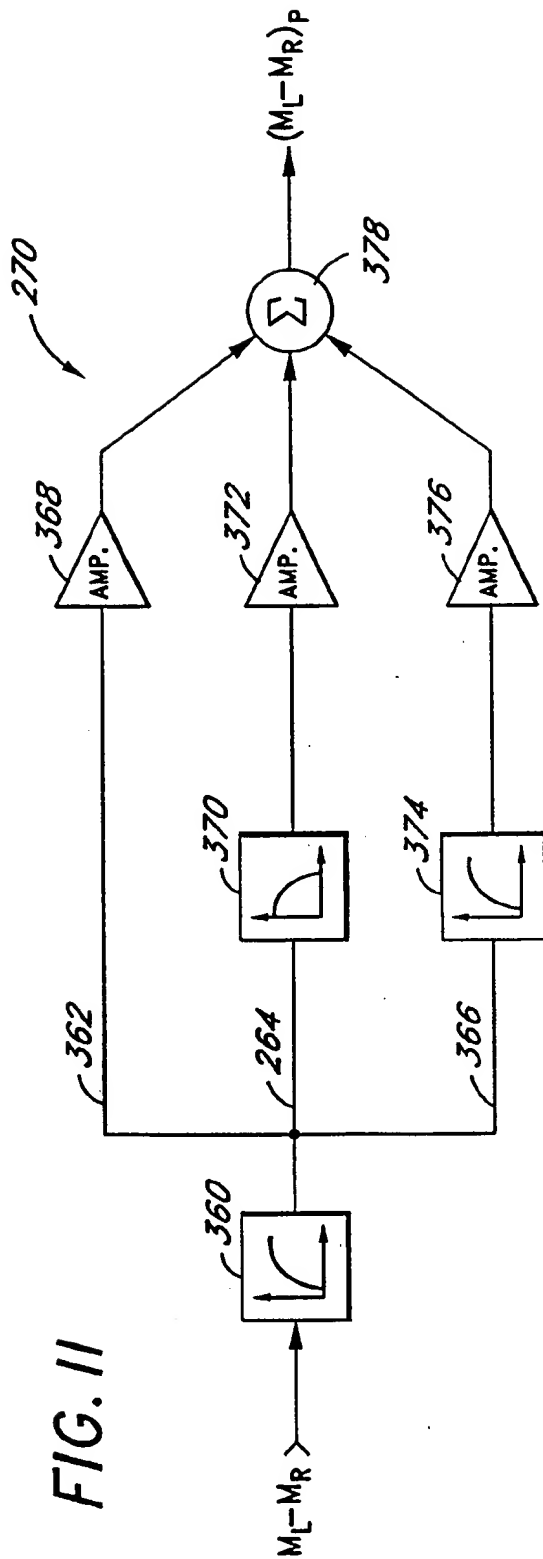
9/10

FIG. 10





10/10



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 97/19825

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 H04S3/00 H04S1/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04S

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
------------	--	-----------------------

A	<p>EP 0 637 191 A (VICTOR COMPANY OF JAPAN) 1 February 1995</p> <p>see page 2, line 22-33 see page 2, line 54 - page 4, line 3 see page 4, line 51 - page 7, line 25 see page 8, line 12 - page 15, line 9</p> <p style="text-align: center;">--- -/--</p>	<p>1-5, 10-20, 29-33, 35, 37-39, 41-44, 46-50</p>
---	--	---

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

3 March 1998

Date of mailing of the international search report

10/03/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Zanti, P

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 97/19825

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>WO 96 34509 A (SRS LABS.) 31 October 1996</p> <p>see page 1, line 4-7  see page 3, line 3 - page 4, line 2  see page 4, line 18 - page 12, line 20  ---</p>	<p>1,6-12,  19-30,  36-40,  43,49,50</p>
A	<p>EP 0 367 569 A (TOSHIBA) 9 May 1990</p> <p>see column 1, line 1-7  see column 4, line 33 - column 19, line 15  ---</p>	<p>1-4,  10-13,  15-17,  19-25,  29-39,  41-43,  46,47,  49,50</p>
A	<p>WO 94 16538 A (DESPER PRODUCTS) 21 July 1994</p> <p>see page 1, line 4-15  see page 5, line 28 - page 9, line 3  see page 11, line 3 - page 12, line 12  see page 13, line 1-37  see page 14, line 8 - page 26, line 28  see page 36, line 13 - page 37, line 29  see figures 7A,7B  -----</p>	<p>1-13,  15-17,  19-43,  45-47,  49,50</p>

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 97/19825

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 637191 A	01-02-95	JP 8051698 A	20-02-96
		US 5579396 A	26-11-96
		JP 7095697 A	07-04-95
WO 9634509 A	31-10-96	US 5661808 A	26-08-97
		AU 5578496 A	18-11-96
		EP 0823189 A	11-02-98
EP 367569 A	09-05-90	JP 2121500 A	09-05-90
		JP 2522529 B	07-08-96
		DE 68927036 D	02-10-96
		DE 68927036 T	06-02-97
		US 5065432 A	12-11-91
WO 9416538 A	21-07-94	WO 9416537 A	21-07-94
		AU 3427393 A	15-08-94
		AU 6081194 A	15-08-94
		CA 2153062 A	21-07-94
		CN 1091889 A	07-09-94
		EP 0677235 A	18-10-95
		JP 8509104 T	24-09-96